

DOCUMENT RESUME

ED 235 633

EC 160 490

AUTHOR Sigelman, Carol K.; And Others
TITLE Communicating with Mentally Retarded Persons: Asking Questions and Getting Answers.
INSTITUTION Texas Tech Univ., Lubbock. Research and Training Center in Mental Retardation.
SPONS AGENCY National Inst. of Handicapped Research (ED), Washington, DC.
PUB DATE 83
GRANT NIHR-16-P-56819/6
NOTE 265p.
AVAILABLE FROM Texas Tech University, Research Training Center in Mental Retardation, Box 4510, Lubbock, TX 79409 (\$15.00).
PUB TYPE Journal Articles (080) -- Reports - Research/Technical (143)
EDRS PRICE MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS Attitudes; *Communication Skills; *Interviews; Language Acquisition; *Mental Retardation; *Questioning Techniques; Research Methodology

ABSTRACT

The study was undertaken as part of an effort to poll mentally retarded persons about their needs, attitudes, and circumstances in order to form an information base for policymaking. Five different studies were conducted to determine the ability of mentally retarded people to understand questions and use speech to convey facts and opinions. Each study involved interviewing mentally retarded children or adults as well as a "significant other." In chapter two, a review of the literature cites research implications for interviewing mentally retarded people. The five study designs are described in the next chapter, with attention to measuring responsiveness and analyzing reliability and validity. The fourth chapter reviews the logistical aspects of the interview approach and offers suggestions for such aspects as obtaining consent, scheduling, and using specific questioning approaches. Chapter five focuses on responsiveness and discusses the difficulties interviewees had in providing appropriate responses. Chapter six evaluates the reliability of interviewee responses over time, while chapter seven assesses the validity of responses (using correct-incorrect questions and agreement with parents/attendants as measures of validity). In chapter eight, direct comparisons of alternative questioning strategies are presented. Following a summary chapter, appendices provide sample interview schedules, suggestions for accessing mentally retarded persons for interviews, and sample explanation/consent forms. (CL)

* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *

**U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)**

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS
MATERIAL IN MICROFICHE ONLY
HAS BEEN GRANTED BY

Gerard J. Bensberg

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC)."

EC160490

COMMUNICATING WITH MENTALLY RETARDED PERSONS:

ASKING QUESTIONS AND GETTING ANSWERS

**Carol K. Sigelman
Carol J. Schoenrock
Edward C. Budd
Jane L. Winer
Cynthia L. Spanhel
Paulette W. Martin
Sherrilyn Hromas
Gerard J. Bensberg**

**Research and Training Center in Mental Retardation
Texas Tech University**

Typist:
Sandra Minor

(c) [REDACTED] 1983

All rights reserved

May not be reproduced in whole or part without written permission of publisher

Research and Training Center in Mental Retardation
Box 4510
Texas Tech University
Lubbock, Texas 79409

This publication was partly supported by grant number 16-P-56819/6 from the National Institute of Handicapped Research, U.S. Department of Education. Grantees undertaking such projects under government sponsorship are encouraged to express freely their judgment in professional and technical matters. Points of view or opinions do not, therefore, necessarily represent official position or policy.

Printed by Texas Tech Press

CONTENTS

PREFACE i.

INTRODUCTION 1.1

The Nature of the Study 1.2

Organization of the Report 1.3

LESSONS FROM THE LITERATURE 2.1

Language and Communication Among Retarded Persons 2.2

Response Effects in Survey Research 2.23

Implications for Interviewing Retarded Persons 2.36

METHODS OF THE STUDY 3.1

Design of Interview Schedules 3.1

The Five Study Samples 3.2

Access to Subjects 3.5

Measuring Responsiveness 3.6

Analyzing Reliability and Validity 3.7

INTERVIEWING RETARDED PERSONS: LESSONS FROM EXPERIENCE 4.1

Designing Interviewing Schedules for the Retarded 4.1

Pilot Testing 4.3

Interviewer Training 4.4

Locating Interviewees and Scheduling Interviews 4.5

Obtaining Consent 4.5

Getting Ready to Interview 4.7

Relating to the Interviewee 4.8

Questioning Approaches 4.9

Closing the Interview 4.14

Recording Interview Data 4.14

Coding and Preparation for Analysis 4.15

Conclusion 4.17

TO WHAT EXTENT CAN MENTALLY RETARDED PERSONS RESPOND TO INTERVIEW QUESTIONS? 5.1

Specific Types of Responses 5.1

The Stability of Responsiveness Scores 5.3

Responsiveness and IQ 5.5

Relationships of Responsiveness to Other Personal Characteristics 5.10

Responsiveness and Question Type 5.10

Effects on Responsiveness of Asking Questions Twice 5.24

Summary and Implications 5.25

TO WHAT EXTENT ARE ANSWERS RELIABLE OVER TIME? 6.1

Reliability Over Time-Institutionalized Children 6.2

Reliability Over Time-Institutionalized Adults 6.2

Individual Correlates of Response Reliability 6.9

Summary and Conclusions 6.11

TO WHAT EXTENT ARE ANSWERS VALID? 7.1

Community Children 7.2

Additional Samples 7.15

Correct-Incorrect Questions as a Measure of Response Validity 7.18

Agreement with Parents and Attendants as an Individual Characteristic 7.23

Conclusions 7.30

WHAT INTERVIEWING APPROACHES ARE MOST USEFUL? 8.1

Question Wording 8.4

Multiple Choice Questions 8.12

Yes-no vs. Open-ended Questions 8.22

Yes-no vs. Either-or Questions 8.25

The Use of Pictures with Yes-no Questions 8.30

The Limits of Acquiescence 8.32

The Use of Pictures with Either-or Questions 8.35

Conclusions 8.40

CONCLUSIONS AND IMPLICATIONS 9.1

Summary of the Study 9.1

How to Interview Mentally Retarded Persons 9.8

Implications of the Study for the Mental Retardation Field 9.12

APPENDIX A A.A

APPENDIX B B.B

APPENDIX C C.C

REFERENCES R.R

Reference Notes R.1

Publications Based on Study R.3

References R.5

PREFACE

This book culminates a broad inquiry regarding what can be expected when questions are asked of mentally retarded persons. It started with a suggestion by the President's Committee on Mental Retardation that it might be useful to conduct a periodic national polling of retarded citizens of all ages and circumstances as a means of gathering input for policy making purposes. The Committee had already commissioned some work on defining the mentally retarded population and constituting a representative sample of that population. It then encouraged the Texas Tech University Research and Training Center in Mental Retardation to propose a study in which the feasibility of interviewing retarded persons would be explored. The study was intended to investigate the ability of retarded persons to answer questions and the reliability and validity of their answers; to compare their perspectives with those of the nonretarded individuals (e.g., parents) who often speak for them; and to

obtain information about the needs, circumstances, and attitudes of retarded persons as well as guidelines for constructing interview schedules in the future.

The proposal was funded by the Rehabilitation Services Administration with the idea that the Center would continue to interact with PCMR in charting directions. This indeed happened, and the President's Committee had a great deal of influence on the subject matter to be included in interview schedules as well as the kinds of research questions to be addressed. We know we did not answer or even address all of the issues that concerned PCMR, but we are confident that this report will be of value to them as well as to a wider range of professionals and researchers in the developmental disabilities field.

The authors have many others to thank. First, Gerard J. Bensberg, Director of the Research and Training Center, was indispensable in getting the project off the ground and supportive

throughout the study. Syng Nam Yough and Cheryl Smith contributed to data analysis. Staff in the following facilities and agencies deserve much thanks for their cooperation in making children and adults available to us for interviewing: the Lubbock State School, the Lubbock Public Schools, particularly the Special Education Department and the Ballenger School; Arkansas Mental Retardation-Developmental Disabilities Services, especially the Arkansas Children's Colony at Conway; and the Austin Association for Retarded Citizens, especially the citizen advocacy program and its staff. Research of this nature always depends on the good will of agency and facility staff, from the top down, and in this project that good will was plentiful. Finally, we reserve special thanks for the parents and attendants, and most especially the retarded people, who actually participated in our interviews. They knew, indeed they were explicitly told, that participating would not

make a difference in their lives, would not help them with the problems that some of them faced. Yet they willingly gave it their best, often enjoying the chance to talk about themselves and to be listened to carefully. All we can say by way of thanks is that we learned more than we ever hoped to learn about retarded people, as well as about how we "non-retarded" people can communicate more effectively with them. Our hope is that this report, even though it uncovers many problems in getting meaningful answers from retarded persons will encourage others to give them a chance to speak for themselves. We will show that good intentions are not enough, that sophistication in phrasing and structuring questions and skill in interacting with people are needed if the attempt is to be fruitful. But as the people we talked to have shown us, there is much to be gained if we try as hard as they do to make the communication process work.

INTRODUCTION

"If more people do not take a more active interest in the life of the mentally retarded then we will not have any place to go except in an institutional living and that is not fair because we are God's creatures, to be like you are, considered as normal." Unsolicited letter cited by Baker, Seltzer and Seltzer (1977).

The mentally retarded adult who wrote these lines was acting as an effective advocate for community residential alternatives to institutions and for the philosophy of normalization. As anyone who has worked with mentally retarded people knows, many of them tell us exactly what is on their minds (whether we want to know or not). Mentally retarded people, like all people, have knowledge of what is going on in their lives, their likes and dislikes, and their hopes for the future. Historically, however, their voices have not been heard. Certainly their perspectives have rarely been solicited by

professionals working with them or researchers studying them. To be sure, there are exceptions. Occasionally, first-hand accounts of the experiences of mentally retarded persons appear in print (see Stanovich and Stanovich, 1979, for a bibliography). Occasionally "we" even solicit their opinions, as illustrated by Nirje's (1972) report of a Swedish conference of mentally retarded adults which yielded a set of provocative policy recommendations. To cite a notable example from the research literature, the deinstitutionalization study by Gollay, Freedman, Wyngaarden, and Kurtz (1978) not only involved interviews with deinstitutionalized persons but also included a screening and critique of proposed interview materials by a group of retarded consumers. Nonetheless, the predominant pattern has been for "us" to speak for "them," prompting Stanovich and Stanovich (1979) to advocate letting "them" talk about us for a change" (p. 83).

In part, neglect of the perspectives of mentally retarded consumers has been a result of the models of retardation that have guided the field. When mentally retarded persons are viewed as subhuman, as children, or as menaces to society, it simply follows that their opinions are irrelevant and that "we" are the ones who know what is and what should be happening in their lives. If, instead, we accept the concept that they are "God's creatures, to be like you are, considered as normal," we are motivated to take them seriously as human beings with important things to communicate. Precisely this concept is gradually infiltrating the field of mental retardation today. The philosophy of normalization has had its impact, as has the broader movement in society to guarantee the civil rights of handicapped persons. Groups of handicapped persons, serving as their own advocates, have had a marked impact in seeing that these changes have come about. The mentally retarded, though rarely speaking for themselves, have been represented by consumer organizations, parents, and other representatives.

One of the major products of the handicapped consumer movement has been the stipulation in a variety of legislation that handicapped clients or students be involved in making those decisions which affect them. Thus, rehabilitation agencies must insure that a client has agreed to an individual rehabilitation plan, mental retardation facilities and agencies must involve clients in the development of habilitation plans, and, under appropriate circumstances, schools must involve not only parents but children themselves in the development of individual education programs. These requirements make it mandatory that human services professionals begin to consider the

perspectives and preferences of mentally retarded persons.

Service providers, counselors, and teachers are not the only ones who have a stake in obtaining information from mentally retarded consumers. Researchers, especially those studying deinstitutionalized and service delivery systems, depend on information gathered from mentally retarded persons, and indeed increasingly express the belief that such persons should be allowed to speak for themselves, since it is their experiences which are under study (e.g., Gollay, Freedman, Wyngaarden, & Kurtz, 1978). Virtually anyone who is involved with the mentally retarded has some need to understand the perspectives of mentally retarded persons and gather information from them.

But how? How does one hear from people whose verbal skills, as a function of the very nature of mental retardation, are limited? It is with this question that the present report concerns itself.

The Nature of the Study

The research was inspired by the President's Committee on Mental Retardation. PCMR became intrigued by the idea of conducting a periodic poll of a representative national sample of mentally retarded persons. The proposed survey, to be used with persons of all ages and living situations, was viewed as a means of collecting information about the circumstances, needs, and attitudes of mentally retarded individuals which could be used to shape national policy. The Research and Training Center in Mental Retardation at Texas Tech was approached with the notion of conducting a study of the feasibility of such a polling, and a proposal was then written and funded by the Rehabilitation Services

Administration. It became clear that PCMR was interested in a wide range of feasibility issues: how to construct a representative sample, how to access interviewees, what to ask, how to ask it, and what to do with the resulting information. In our own thinking about the project, we decided that the first and most important task was to determine what can be expected of mentally retarded persons in an interview situation; *whether they can give answers, and whether their answers are reliable and valid sources of information about their circumstances, needs, and attitudes.* If we assume that mentally retarded persons should be given opportunities to communicate about their own lives and needs, it is critical to explore the methodological issues involved in obtaining such information.

In an important sense, then, the study became an exploration of the communication skills of the mentally retarded; not their language development, but their ability to understand questions and use speech to accurately convey facts and opinions. However, our intent was not just to explore deficiencies in communication skills among the mentally retarded. Communication is a two-way process, and the ability of mentally retarded persons to respond to a question might depend heavily on the form, clarity, and salience of the question. Thus we sought not only to understand the limits of retarded people's abilities to participate in interviews but also to identify more and less effective ways of asking questions of them.

The four guiding questions of the project can be stated as follows:

1. To what extent can retarded persons respond to questions in an appropriate fashion, and what factors affect their responsiveness?

2. How reliable are such responses, in the sense of being consistent over short periods of time?
3. How valid are such responses, in the sense of being free of systematic biases and agreeing with information provided by parents or caretakers or documented in records?
4. What types of questions appear to optimize responsiveness, reliability, and validity?

Organization of the Report

The body of this report is based on data obtained from five different interviewing studies. Each study involved interviewing a group of mentally retarded children or adults about their circumstances, needs, and attitudes, and also interviewing a "significant other" such as a parent or attendant to obtain a second perspective on each client. In Chapter 2, we briefly review the literature relevant to the project, considering both the literature on the communication skills of mentally retarded persons and the literature on the potential for biased responses in survey research with the general population. In Chapter 3, we describe the design of our study, the samples interviewed, and the measurement procedures used so that this methodological information can be borne in mind in reading later chapters. In Chapter 4, the logistical challenges in conducting interviews are considered, and we offer suggestions based on our experiences for carrying out interviews with retarded persons.

Because the five studies were very similar in basic purpose and design, we do not report the findings study by study. Instead, the remainder of

the report is organized around the four guiding questions posed above. In Chapter 5, we draw from the various studies to examine responsiveness, or the extent to which mentally retarded persons are able to answer questions. We describe the kinds of difficulties interviewees had in providing an appropriate answer; for example, in saying yes or no to a yes-no question or naming something in response to an open-ended question. We look at responsiveness as an individual characteristic, describing how low in the IQ range one can go and expect mentally retarded persons to be able to respond to simple questions, and we relate responsiveness to other factors such as the type of question asked (yes-no versus either-or versus open-ended, and so on). Starting with Chapter 6, we move beyond the issue of whether or not various questions elicit appropriate answers and ask whether or not the responses obtained are useful. Chapter 6 focuses on the issue of reliability by analyzing data obtained in two of our studies in which institutionalized persons were asked the same questions twice, a week apart. If, in an interviewing study, the answers given by a person change drastically from week to week, the information obtained is of little use. Again we examine factors (such as question format) that affect the extent to which answers are stable from week to week. In Chapter 7, we turn to the validity issue, primarily by examining the extent to which answers given by retarded persons and answers given by "significant others" agree. Although we do not assume that significant others represent "the whole truth," we are concerned with the extent to which pictures of a group provided by clients and their parents or attendants differ.

In Chapter 8, we describe direct comparisons of alternative ways of requesting the same information. Throughout the five studies, a number of different comparisons of this type were made. In some cases, we compared alternative question formats (e.g., yes-no questions versus either-or questions on the same topic), while in other cases we explored the effects of alternative question wordings or alternative probing techniques. In each case, we compared the questioning strategies at issue with reference to three criteria: responsiveness, absence of systematic response biases, and validity as measured by agreement with parents or attendants. Thus, we first asked which alternative yields appropriate responses from a higher proportion of the population, assuming that the more useful questioning approach is the one to which most clients can respond. Then we compared the two alternatives in terms of the extent to which systematic biases, revealed in comparison of a client's responses to one format with his or her responses to the other, invalidate responses. Finally, we determined which question is associated with higher agreement with the responses given by parents or attendants, on the assumption that high agreement with significant others is an additional indicator of validity of responses. Of any two approaches to soliciting the same information, then, the more useful of the two is the approach which optimizes responsiveness, does not appear to be associated with any systematic response biases, and yields answers that concur with those provided by significant others. Chapter 8 thus has many practical implications for anyone seeking improved methods of obtaining information from mentally retarded persons.

Finally, in Chapter 9 we attempt to summarize the major findings of the study and their implications for

training mentally retarded persons in communication skills and for seeking information from them as a service provider, counselor, researcher, or simply interested party. Some readers may want to read Chapter 9

first to get an overview of the study and its findings and then go back to chapters of special interest. Finally, we have provided a variety of appendix material of particular interest to those planning interview studies.

LESSONS FROM THE LITERATURE

As in any research project, we began our inquiry with reviews of previous research, hoping to find a body of knowledge on which to base our own study. We did not find a literature on how best to ask questions of retarded persons. As far as we can determine, ours is the first such study. However, we searched for clues in two related bodies of literature; the literature on language and communication skills of the mentally retarded, and the literature on methodological issues in survey research with the general population. This chapter overviews what we found most pertinent to our research questions in these two very different bodies of literature.

Information about the communication skills of mentally retarded persons is extensive, as deficits in language and communication are part of the very nature of mental retardation. As experienced mental retardation researchers, we of course knew that the limited verbal skills of many retarded persons would limit the kinds of

questions that could be asked of them. We knew that verbal interviewing would be all but impossible with profoundly retarded persons, but were not as sure how much could be expected of those at higher levels of retardation. We hoped to gather information about whether we could expect of a retarded person essentially what could be expected of a normal child of the same mental age, or whether some of the language and communication difficulties of retarded persons are not so much a problem of developmental delay as a matter of qualitative difference. For example, if retarded persons have been found to have serious difficulties in understanding quantitative concepts and terms, beyond those that could be predicted on the basis of mental age, we would want to avoid their use in interviews.

Reviewing this vast literature was difficult, for participating in an interview is an activity that calls on all of a person's language capacities. Receptive skills are needed

to understand the question being asked; expressive skills to give an answer. Moreover, reception and expression, or comprehension and reproduction, must be applied to all of the three subsystems of language: sound, meaning, or semantics, and grammar or syntax. Thus, we had to be concerned with hearing deficits and speech or articulatory problems among retarded persons; we had to explore their vocabulary and comprehension of various concepts; and we had to consider their ability to understand the structure of questions posed to them as well as to structure answers.

The selective literature review we present here considers in turn the sound, meaning, and grammatical systems of language and suggests what can be expected of retarded persons in interview situations. However, this body of literature did little to tell us whether the answers of retarded persons can be expected to provide meaningful insight into their lives. We can ask a simple yes-no question, for example, and a retarded person may appear to understand the question and answer "yes." But is that "yes" a reflection of reality? Would the person's parent or someone else who knows the person well concur? These are issues which have not been addressed in the mental retardation field; yet they are critical if we are concerned not just with linguistic competencies but with the use of those competencies for a purpose: to communicate accurately one's needs, circumstances, and attitudes.

Fortunately, we had another body of literature to turn to for clues concerning the issue of response validity. Researchers who use survey research methods with the general population have long been concerned with threats to response validity. They have repeatedly found, for example, that many people

give biased answers to put themselves in a favorable light and that how one structures and phrases a question may influence what answers are given. We assumed that if certain biasing effects have been uncovered in survey research with the general population, there is no reason to believe that they will not surface in interviews with retarded persons. On the contrary, as we will show, there is reason to believe that threats to response validity will be even more acute in interviews of retarded persons than they are in interviews of nonretarded persons. Thus this chapter's second section will overview what survey researchers have discovered about response bias.

Language and Communication Among Retarded Persons

Speech and communication problems are generally viewed as characteristic of mentally retarded persons (Blount, 1968; Karlin & Strazzula, 1952; Keane, 1972), and have frequently been used as important diagnostic indicators of mental retardation (Blount, 1969; Matthews, 1971). Unfortunately, research has provided little consistent normative data on the general level of linguistic competence to be expected in retarded persons of a particular intelligence level (Matthews, 1971). This lack of general expectations about language competencies can be attributed to the varied approaches used to study language development in the retarded (Yoder & Miller, 1972). Studies are difficult to compare directly due to methodological differences, variations in the populations studied, and differences in the specific types of language behavior measured. However, those who have surveyed the literature of language development of the mentally retarded have been able to draw some broad conclusions.

LANGUAGE AND INTELLECTUAL FUNCTIONING

The retarded child typically acquires language and speech considerably later than the nonretarded child (Baroff, 1974; Cromer, 1974; Matthews, 1971; Yoder & Miller, 1972). While many developmental sequences are delayed for retarded children, the areas of speech and higher intellectual functions show the greatest amount of delay (Karlin & Strazzula, 1952). For example, Schlanger (1953a) reported that the average age of onset of speech in a group of institutionalized retarded children was slightly over three years, while nonretarded children begin speaking about the first year (McCarthy, 1954). However, even though retarded children are reported to lag behind their nonretarded peers on such measures as sentence length, sentence complexity, discrimination of speech sounds, and percentage of nouns used (Spreen, 1965), most authors agree that the course of language development generally parallels that of normal children. That is, it is not qualitatively different (Lackner, 1976; Lenneberg, Nichols, & Rosenberg, 1964; Yoder & Miller, 1972).

Baroff (1974) has attempted to provide some rough expectations regarding the linguistic competence of retarded children and adults as a function of mental age and degree of retardation. As shown in the summary of these norms (Table 2.1), Baroff suggests that the major language deficit for retarded persons involves expressive language abilities. As has been found true of normal children, retarded persons are usually able to understand more speech than they are able to produce. Therefore, a general strategy for interviewing the retarded might be to devise questions which minimize productive demands (e.g., through simple yes-no questions, or forced-

choice questions, or even questions that require only pointing).

Baroff's norms for retarded children also suggest that the order of language acquisition follows the developmental sequence reported for nonretarded children (McCarthy, 1965). For our purposes, the most useful information that emerges from Baroff's data is that the ability to respond to questions is not evident prior to mental age four or five. Therefore, efforts directed at interviewing retarded persons below this mental age may be inadvisable and of limited value.

Baroff also emphasizes that speech problems and language delays are more pronounced at the lower levels of intellectual functioning, and that speech development in the retarded child is closely associated with mental age. For the most part, research on the language development of the retarded supports these observations (Karlin & Strazzula, 1952; Lyle, 1961b; Sirkin & Lyons, 1941).

As a general rule, then, the higher the level of intelligence, the better the language ability of the retarded individual (Keane, 1972). Many early studies indicated that speech and language measures are positively correlated with intellectual level in retarded populations, even though these correlations are often of only a moderate degree. For example, Spradlin (1963) reviewed four studies which correlated articulation and intelligence measures for institutional populations and found that the coefficients ranged from .41 to .58. Schlanger (1958) found that auditory word discrimination correlated .59 with mental age, and other studies have reported correlations ranging from .42 to .76 between vocabulary size and intelligence (Condell, 1959; Mein, 1962; Sloan & Bensberg, 1954). By contrast, chronological age has little or no

Table 2.1: Summary of language norms for retarded children and adults as a function of mental age, chronological age, and degree of retardation from Baroff (1974)

Mental Age	Chronological Age	Degree of Retardation	Language Behavior
1 yr	4+ years	Profound	Expressive: Imitates sounds. Laughs or smiles in response. May say "mama" or "dada". At a pre-speech level: crying, vocalization, and gestures Receptive: Some understanding of gestures and very familiar words (e.g., "no").
1-2 yrs	3-6 years 8+ years	Severe	Expressive: May use one or two words, but pre-speech forms continue to predominate.
2-3 yrs	3½-5 years 5-7 years 6-9 years 10+ years	Mild Moderate Severe Profound	Expressive: May have considerable speaking vocabulary though speech will be particularly impaired in children with more than mild retardation. If nonverbal, as in severely and profoundly retarded, there may be use of gestures to communicate. Receptive: Understand simple verbal communications, (e.g., following directions, responding to questions).
4-5 yrs	6-9 years 10-12 years 12-15 years 16+ years	Mild Moderate Severe Profound	Expressive: May have considerable speaking vocabulary though speech will be particularly impaired in children with more than mild retardation. If nonverbal, as in severely and profoundly retarded, there may be use of gestures to communicate. Receptive: Understands simple verbal communications, (e.g., following directions, responding to questions).
5-6 yrs	8-10 years 11-13 years 13-15 years	Mild Moderate Severe	Expressive: Mildly retarded child may be using fairly normal sentence structure and have speech of good intelligibility. At more severe levels of retardation, language may be at the phrase or single-word level and also be indistinct in pronunciation.
6-7 yrs	10-12 years 14+ years	Mild Moderate	Expressive: May be able to carry on a simple conversation and use complex sentences (more true of mildly than moderately retarded children).
7-11 yrs	13-15 years	Mild	Expressive: Essentially normal fluency though pronunciation problems persist.

Table 2.1: continued

Mental Age	IQ Level	Degree of Retardation	Language Behavior
8-11 yrs	52-68	Mild	Language is adequate for ordinary purpose of communication at both expressive and receptive levels.
6-8 yrs	36-51	Moderate	Language is generally functional for purposes of communication, but intelligibility of speech may be much impaired.
4-6 yrs	20-35	Severe	Understanding language is likely to be much better than the ability to express it. Speech may be very poorly articulated and difficult to understand.
0-4 yrs	0-19	Profound	Ability to understand is likely to far exceed ability to speak. There may be little or no speech. Capable of following simple directions.

relationship to linguistic ability in the retarded child (Bangs, 1942; Schlanger, 1953b). The relationship between intelligence and language generally holds for persons with a mental age of eight years or less, at which point the degree of correlation diminished (Spreen, 1965). It is interesting to note that Goda and Griffith (1962) have reported similar results for nonretarded children. They contend that speech development is directly related to intellectual growth in normal children until age eight. After this point, further development is minimal and consists mainly of mastering more complex sentence structures and gradually increasing vocabulary size. Therefore, prior to mental age eight, one would expect to find language acquisition and proficiency closely related to intellectual level in both retarded and nonretarded populations. The language development of the retarded child would be very similar to that of the nonretarded child of an equivalent mental age. Consequently, retarded persons below mental age eight would not have mastered the fundamental aspects of language and would evidence more immature speech forms than nonretarded persons of the same chronological age.

While most authors report that speech impairments and language deficits are more numerous and severe in retarded populations, it is generally accepted that no specific types of language disorders are found in the speech of the retarded that are not also found in nonretarded populations (Keane, 1972; Matthews, 1971; Spreen, 1965). Similarly, Lenneberg and his associates (1967; Lenneberg, Nichols, & Rosenberg, 1964) have concluded that no structures are found in the language of the mentally retarded that do not occur in the course of development for nonretarded children. These authors contend that differences observed in the speech and

language patterns of the retarded are primarily quantitative in nature and do not represent qualitative language differences.

In contrast to previous research, however, Lenneberg et al. (1964) reported that chronological age was a better predictor of language achievement than intelligence in a three-year study of Down's Syndrome children whose ages ranged from 3 to 22. They also concluded that language acquisition in Down's Syndrome children depends more on physiological maturation than on cognitive development. Lenneberg's findings highlight two important considerations. First, etiology plays a role in determining the course of speech development and competence in the retarded (Cromer, 1974; Spreen, 1965; Webb & Kinde, 1967). Therefore, it would seem necessary to distinguish type of mental retardation as well as degree of retardation when assessing the linguistic ability of the retarded. Secondly, research has suggested that factors other than cognitive ability may influence language acquisition, and therefore we must recognize how such factors affect the language behavior of the retarded.

LANGUAGE AND ETIOLOGY

Early research designed to identify specific language deficits among the various diagnostic categories of mental retardation has typically compared Down's Syndrome or brain-damaged groups with other groups of mentally retarded persons. In general, brain-damaged retarded children show the greatest amount of unevenness in their expressive and receptive language abilities (Rigrodsky & Goddard, 1951). For example, when compared to non-brain-damaged groups, they are inferior on auditory discrimination and perception tasks (Spreen, 1965; Schlanger, 1958), and are inferior in making verbal associations, integrating verbal concepts, and using

visual-motor perceptual abilities in speech production (Gallagher, 1957). On the other hand, the brain-damaged as a group typically have larger vocabularies, use more advanced modes of verbal expression, give more detailed and complete definitions for words (Bijou & Werner, 1944, 1945), and are better on verbal imitation and speech production tasks (Gallagher, 1957).

In contrast, Down's Syndrome children as a whole are consistently poorer on most language measures when compared to other groups of retarded children (Lyle, 1960c, 1961a, b). Mein (1961) has noted that Down's Syndrome children decrease their use of nouns at a later age than do most other groups of retarded children, which has been interpreted as representing a more concrete and immature style of thinking. Speech defects are also reported to occur more frequently among Down's Syndrome children (Schlanger & Gottsleben, 1957). In particular, several authors have noted a greater incidence of stuttering among Down's Syndrome groups (Gottsleben, 1955; Schaeffer & Shearer, 1968; Schlanger, 1953a). Finally, Bilovsky and Share (1965) suggest that a deficit in auditory-sequential memory is a major factor in the language retardation of Down's Syndrome children.

Many of these findings have been supported and extended in a recent investigation by Rohr and Burr (1978). They compared four diagnostic groups of retarded subjects (Down's Syndrome, biological brain damaged, environmentally caused, and unknown etiology) according to their performance on the Illinois Test of Psycholinguistic Abilities. All four groups exhibited a similar pattern of scores on the subtests, doing best in verbal expression and worst in auditory memory. However, it was found that the Down's Syndrome group consistently scored lower than the other three groups on

all auditory-verbal subtests except verbal expression. In contrast, the environmentally-caused retardation group scored higher on all subtests than did the other diagnostic groups, although the differences were not always significant. The scores for the other two diagnostic groups were more variable, but generally fell between the extremes defined by the Down's Syndrome group and the environmentally-caused retarded group. The authors suggest that an auditory memory deficit contributes to the language problems of all retarded groups, with Down's Syndrome children simply being affected the most rather than being unique as Bilovsky and Share (1965) implied. Therefore, etiology appears to have an overall influence on the degree of auditory-verbal deficiencies in the mentally retarded, but no specific pattern of language disability can be linked to a specific type of retardation.

LANGUAGE AND OTHER FACTORS

Matthews (1971) suggests that there may be many other explanations for poor speech development and language competence besides low intellectual functioning. For example, emotional disturbances, hearing loss, and lack of verbal stimulation may delay language acquisition or result in abnormal speech patterns.

Environmental deprivation is frequently cited as producing negative effects on the language development of children, and the adverse effects of institutionalization on language development has been studied extensively in both retarded and nonretarded populations (Spradlin, 1963; Spreen, 1965; Spitz, 1945, 1946; Skeels, Udegraff, Wellman, & Williams, 1938). Institutional living has been associated with lower scores on vocabulary measures in both normal (Skeels et al. 1938) and retarded children (Little & Williams, 1937). Schlanger (1954) reported

that institutionalized retarded children produce shorter sentences, while Sievers and Essa (1961) found that institutionalized children scored lower on the Developmental Language Facilities Test than did a matched group of noninstitutionalized children. Finally, Lyle (1959, 1960a, b, 1961b) has concluded from a series of studies on the effects of institutionalization that noninstitutionalized retarded children are superior on overall language ability and that the negative effects of institutionalization may be greater at certain points in the language development sequence. These authors point to the effects of family separation, emotional deprivation, lack of good speech models, and lack of speech motivation as the causes of these deficits.

More recent research has sought to determine if some language functions are more affected by institutionalization than others are. For example, Montague, Hutchinson, and Matson (1975) analyzed psychological and sociological language content categories and found little difference between language samples of institutionalized and noninstitutionalized retarded children on these measures. They suggest that institutionalization may produce the same negative effects on language content as were reported previously for measures of language structure and vocabulary.

More recently, McNutt and Leri (1979) examined both the structure of spoken language and specific auditory-verbal abilities as measured by the Illinois Test of Psycholinguistic Abilities among institutionalized and noninstitutionalized retarded children.

Their findings supported the suggestion of Montague et al. (1975) that institutionalization may have selective effects on the language abilities of the retarded. In particular, they found that their institutionalized groups had significantly poorer scores on auditory reception, verbal expression, and auditory closure. However, in contrast to previous research, they found no differences between the two groups on measures of language structure. Overall, then, when differences are found between groups of institutionalized and noninstitutionalized retarded persons, they generally favor those who do not live in institutions, but more markedly in some areas of competence than in others.

In addition to examining the effects of institutionalization, current research has also looked at differences in the early mother-child verbal interactions of retarded and nonretarded infants. Verbal interactions between retarded children and their mothers appear to be different from the communication patterns of nonretarded children and their mothers (Buism, Rynders, & Turnure, 1974; Gutmann & Rondal, 1979; Marshall, Hegrenes, & Goldstein, 1973). For example, Buism et al. (1974) found that mothers of Down's Syndrome children appear to be modeling shorter and grammatically less complex verbalizations for their retarded children than mothers of nonretarded children do. Currently, we do not know whether such verbal modifications made by the mothers of retarded children adversely affect their children's language development. Gutmann and Rondal (1979) have suggested that these adaptations may not be functionally inappropriate to the retarded child's language comprehension.

needs; however, they may contribute to the simple language structure generally reported to occur in the speech of the mentally retarded.

In summary, then, speech and communication problems abound among retarded persons, with the major difficulties being evident in the expressive area. Even though language development appears to be delayed in the majority of retarded children, the sequence of development has been found to parallel that of nonretarded children and is not considered to be qualitatively different. A lack of general language expectations prevents us from making accurate predictions as to what can be expected of the retarded person of a particular mental age; however, on the average, language competence increases and speech problems decrease with higher intellectual functioning. In addition, a number of other factors such as environment may contribute to the language problems of the retarded and produce more severe language deficits than would be expected from their level of intellectual functioning. We will now focus on more specific speech and language deficits among the retarded that should be taken into account in interviewing them.

PROBLEMS RELATED TO SPEECH AND HEARING DEFECTS

Research on the prevalence of speech and hearing disorders among the mentally retarded has been reviewed extensively by several authors (Bensberg & Sigelman, 1976; Keane, 1972; Matthews, 1971; Spradlin, 1963; Spreen, 1965). There is substantial evidence that speech and hearing defects are more frequent in mentally retarded populations than in the general pop-

ulation. However, wide variation among prevalence figures is also evident.

Speech Defects

Speech defects among retarded interviewees may make it difficult for them to express themselves and for interviewers to understand them. Spradlin (1963) reviewed 14 studies which surveyed the prevalence of speech defects in mentally retarded children and concluded that the percentage of institutionalized retarded persons having speech defects varies from 57 to 72 percent. Matthews' (1971) review of the literature reported figures which varied from 18 to 94 percent, with median rate of 56 percent. Thirteen of the 18 retarded samples studied were also from institutional settings.

According to Spreen (1965), the prevalence of speech defects varies as a function of intelligence within the retarded population. He estimates that almost 45 percent of those classified as mildly retarded experience speech problems of some kind. This figure increases to 90 percent in the moderately retarded range, and nearly all children in the severely and profoundly retarded ranges demonstrate some kind of communication disorder, which may include lack of speech.

Those authors who have attempted to assess the frequency with which specific speech problems occur among the mentally retarded tend to agree that articulation and voice disorders constitute the largest percentage of speech difficulties (Keane, 1972; Spradlin, 1963). There is evidence to indicate that this is also true in nonretarded populations (Spradlin, 1963) and that, in general, the types of speech problems found in retarded groups are also found in normal

populations (Keane, 1972; Spradlin, 1963; Spreen, 1965).

Stuttering is another speech difficulty which, like articulation and voice disorders, seems to be especially common among the retarded. Keane (1972) rates stuttering as the third most common speech defect of the mentally retarded, but noted that the higher frequency of stuttering in this group is still controversial. For example, Karlin and Strazzula (1952) observed only one case of stuttering out of 50 mentally retarded children, a finding which corresponds with the prevalence figures typically cited for the general population (2%). There seems to be more agreement that a higher than normal incidence of stuttering exists among persons with Down's Syndrome (Shubert, 1966; Schlanger & Gottsleben, 1957; Gottsleben, 1955). In addition, voice disorders such as low husky voices are reported to be especially common among Down's Syndrome children (Benda, 1946; Karlin & Strazzula, 1952; West, Kennedy, & Carr, 1947).

Mentally retarded persons have also demonstrated difficulty in making sound discriminations, which contributes to poor speech production. In one study, Schlanger (1953b) determined that only 19 percent of his sample of retarded children ages 8 to 16 years were able to make sound discriminations at the second grade level of proficiency, even though the average mental age for the sample was 6-8. In a previous investigation, Schlanger (1953a) reported that these same children were also inferior to nonretarded children matched on mental age in auditory memory span for vowels.

The development of speech sounds and speech proficiency in retarded children has been examined in sev-

eral studies. Matthews (1971) cites a 1942 study by Bangs as one of the most systematic studies of typical errors found in the speech of the mentally retarded. Bangs analyzed the types of verbal substitutions, omissions, and additions made by 53 retarded children. He found that mental age was more closely related to speech proficiency as measured by the number of articulation errors ($r = .41$) than either intelligence ($r = .33$) or chronological age ($r = -.11$). In general, the types of articulatory errors made by retarded children were very similar to those produced by nonretarded children of a comparable mental age.

During the same year, Irwin (1942) reported on the speech sound development of ten mentally retarded children ranging from one to five years old over a one-year period. Unlike Bangs, he concluded that speech sound development in his sample of retarded children was qualitatively different from that of normal children, because their level of proficiency approximated that of normal children less than one-year old. However, Irwin's findings may actually indicate that speech sound acquisition was severely delayed in this group rather than being different from that reported for normal children because his sample was still quite young.

A more extensive study was reported by Karlin and Strazzula (1952), using subjects with a wider range of intelligence scores and ages than those studied by Irwin. The authors observed that the majority of children had speech defects which were similar to those found in normal children, but that the defects were more prevalent and often more severe. Similarly, Schlanger's (1953a, 1953b) data also point to a delay of normal

speech development in retarded children, rather than the development of deviant speech patterns, and support Bangs' rather than Irwin's findings. Finally, Goda and Griffith (1962) found that both intelligence and mental age were related to articulation scores in a sample of retarded adolescents ($r = .58$ and $.52$, respectively).

Therefore, when interviewing the mentally retarded one would expect to find the same types of speech defects commonly found in normal children, but with greater frequency and often greater severity. The evidence seems to suggest that speech proficiency is related to the retarded individual's level of intelligence, and one could expect the speech capacity of the retarded person to correspond to that of a nonretarded child of a comparable mental age. Several authors have indicated that articulation and voice disorders are the most common types of speech difficulties among the retarded. This would suggest that interview techniques which make minimal expressive demands on the interviewee would be most effective with mentally retarded persons.

Hearing Disorders

Hearing disorders not only interfere with language development but obviously make for difficulties in understanding questions asked in an interview. Like the literature on speech disorders, the literature on hearing disorders among the mentally retarded suggests wide variation among the estimated percentages of retarded persons with hearing impairment. However, the estimates have consistently exceeded the five percent prevalence figure commonly cited for hearing losses in normal public

school children. Individual studies have reported estimated percentages of hearing problems in retarded populations ranging from two-and-a-half to eighteen times that found in the general population (Birch & Matthews, 1951; Rittmanic, 1959; Schlanger & Gottselben, 1956). Therefore, despite conflicting figures, the consensus is that the occurrence of hearing problems is considerably greater in this population than in nonretarded populations (Bensberg & Sigelman, 1976; Matthews, 1971; Lloyd & Reid, 1967).

One major consideration in determining the prevalence of hearing disorders in any population involves the choice of criteria for defining a hearing loss. Several authors have noted that variation in hearing loss criteria has created problems in comparing prevalence figures (Webb & Kinde, 1967; Kodman, 1958; Lloyd & Reid, 1967). Kodman (1958) reviewed seven studies on the prevalence of hearing problems in the retarded and found that the rate of hearing impairments ranged from 13 to 49 percent. He proposed that the criteria for defining a hearing problem should be standardized to represent a loss of 30 decibels or more in either ear at one or more frequency (range 125 to 8000 cycles per second), on the basis that most hearing losses below this level are not severe enough to interfere with speech reception and language acquisition. According to these criteria, he estimated that the prevalence of hearing losses among the mentally retarded would then average 15 to 20 percent. While Lloyd (1970) agreed with these estimates, he argued that a less stringent criterion for hearing loss is more appropriate.

Most studies which have examined the severity of hearing disorders at various levels of retardation have found almost no relationship between the percentages of hearing problems observed and the level of measured intelligence (Reynolds & Reynolds, 1979; Rigrodsky, Prunty, & Glovsky, 1961; Schlanger & Gottsleben, 1956; and Siegenthaler & Krzywicki, 1959). However, this may be an artifact, since severely and profoundly retarded persons are often untestable by audiometric screening (Hogan, 1973). In addition, hearing disorders are known to increase with age in all populations, and there are several studies which demonstrate this fact in retarded populations (Kodman, 1958; Kodman, Powers, Philip, & Weller, 1958; Schlanger & Gottsleben, 1956). Schlanger and Gottsleben (1956) observed a 25.7 percent hearing loss rate in their under 20 age group, but this percentage increased to 41.4 percent in their over 20 age group. Kodman et al. (1958) used the same hearing loss criteria and obtained more conservative figures. Their young group (mean age = 15.4 years) demonstrated a 19.0 percent rate of hearing loss, while their older group (mean age = 38.7 years) showed a 23.8 percent hearing loss rate.

Even though there is considerable variation among the surveys of hearing problems in the retarded, on the average, hearing impairments seem to occur four to five times more often in retarded children than in normal children. Studies which use less stringent criteria to define a hearing loss and which evaluate older subjects frequently show higher prevalence figures. Finally, although there is no solid research evidence, one must consider the possibility that a hearing loss in a retarded

child is more handicapping than the same loss in a normal child. Lloyd and Cox (1972) have suggested that "the retarded do not demonstrate the same adaptive behavior or capacity to compensate for sensory involvement such as auditory or visual losses as do nonretarded individuals" (p. 22). Possibly then, a hearing handicap of even minor severity could represent a serious disability to the retarded person and decrease his or her ability to learn about the world and communicate with others.

The Semantic System of the Mentally Retarded

Semantics or meaning in the language of retarded persons has implications for how well they can understand the words, concepts, and questions put to them. Studies which compare the vocabularies and lexical usage of retarded children and non-retarded children focus primarily on such things as vocabulary size and diversity, degree of abstraction in word usage, word frequencies, and the relative usage of various parts of speech. In general, most studies indicate that retarded children use significantly more concrete word forms and references than do normal children of a comparable mental age (Yoder & Miller, 1972). For example, Sloan and Cutts (1947) determined that the essential difference between "easy" and "difficult" items on the Stanford-Binet Intelligence Scale for retarded children was the degree of abstraction required to pass an item.

Several early studies examined the retarded person's ability to formulate abstract word definitions. Papinia (1954) demonstrated that as mental age increased, the ability to define words in more abstract terms also improved for retarded persons. At any given mental age level retarded children could acceptably define as

many words as nonretarded children but their definitions were often simpler and more concrete. These findings suggest that retarded children do not differ qualitatively from normal children in terms of defining words, but merely develop the ability to provide more abstract definitions later.

Badt (1958) has suggested that length of institutionalization may be a significant factor influencing the level of abstraction exhibited by retarded persons. Abstraction scores for a group of moderately and mildly retarded state school residents were negatively associated with the number of years each child had spent in the institution ($r = -.61$), while chronological age and mental age were only modestly related to level of abstraction ($r = .34$ and $.24$, respectively). Badt concluded that institutional living interferes with the retarded child's ability to formulate abstract relationships and to manipulate concepts.

The inability of retarded children to express abstractions is also seen in their performance on tests which call for identifying similarities among different objects. Retarded children frequently insist on naming differences rather than abstracting a common feature from a group of items. For example, Griffith and Spitz (1958) found that retarded boys had trouble abstracting a common characteristic from a group of words similar on one dimension, even though they could adequately define the words individually.

Therefore, the evidence seems to indicate that retarded children are less apt than normal children to use abstract concepts. In addition, there is some evidence to suggest that this ability to use language at a more abstract level becomes more

impaired the longer a child remains institutionalized. This line of research, even though it deals with production rather than comprehension, suggests that retarded interviewees may have special difficulty understanding questions which contain abstract words or concepts.

The size and composition of an individual's vocabulary is another important aspect of semantic development. Mein and O'Connor (1960) compared the oral vocabularies of institutionalized retarded persons with those of nonretarded children of a comparable mental age. Retarded persons were found to average nearly one-third more words in their individual vocabularies than nonretarded children. Spreen (1965) reviewed several studies confirming this observation that mentally retarded persons tend to have larger vocabularies than nonretarded children of the same mental age. He attributes their larger vocabulary size to the fact that mentally retarded subjects are typically older and have had more exposure to language than their nonretarded counterparts.

Looking at the kinds of words in the vocabularies of retarded and normal children, Mein and O'Connor found that institutionalized retarded persons share more words in common than do normal children of the same mental age, possibly because institutionalization causes them to have more experiences in common with their peers. Nonretarded children exhibited greater individuality in their vocabulary, despite the fact that their total vocabulary size was smaller. It should be noted that in this study the size of a retarded individual's vocabulary did not seem to be affected by the length of institutionalization ($r = .17$) but was significantly

related to the retarded person's mental age ($r=.72$)

In a study comparing word frequency and word usage of state school residents and nonretarded children, Beir, Starkweather, and Lambert (1969) observed only minor differences in the types of words used by retarded and nonretarded children. For example, retarded children tended to use positive words like "yes" and "O.K." more often (see the discussion of acquiescence that follows). Retarded children also used a greater number of self-references (me, I, my, etc.) and fewer other-references, which is a pattern common in very young children and indicative of social immaturity.

Studying vocabulary usage in terms of common versus uncommon words, Lozar, Wepman & Hass (1972) found that the verbal output of retarded children was very much like that of nonretarded children. However, they also noted that there were considerable differences in the retarded child's usage of different parts of speech. They collected speech samples from institutionalized retarded children and normal children matched on either mental age or chronological age by asking them to tell stories to TAT cards. The retarded children often spent more time describing the cards which led to a higher proportion of nouns, articles, and quantifiers in their speech. On the other hand, the normal children were more likely to tell stories about the pictures, producing more verbs, prepositions, and conjunctions.

Previous research with nonretarded adults has suggested that the earlier a word is estimated to have been acquired by an individual, the faster the word is retrieved from memory (Carroll & White, 1973a, 1973b; Lachman, 1973; Lachman, Schaffer, & Hennrikus, 1974). Therefore Winters and

Brzoska (1975) have suggested that the sequence in which words originally entered semantic memory can be estimated from their relative speech retrieval. Using this procedure, they found that retarded persons have smaller lexical stores than would be expected compared to nonretarded persons of a similar mental age, but that they displayed a similar sequence of acquisition of lexical items. An unexpected finding however, was that the degree of lag varied according to the class of the word, such that retarded persons were most deficient in words normally acquired earlier in the development sequence. This finding, though paradoxical, does support earlier findings that as mental age increases retarded persons show more diversity in their lexical usage (Bartel, Bryen, Keehn, 1973).

In a follow-up to this study, Winters and Cundari (1979) compared the speed of retrieval of words from the lexicon of institutionalized mentally retarded adolescents to the age-of-acquisition estimates derived from nonretarded adults. Words that were estimated to have been acquired the earliest by the normal group were retrieved the fastest by the retarded group. The authors interpreted these results as indicating that the verbal memory process of retarded and nonretarded persons are quite similar, despite the differences in lexical store size found by Winters and Brzoska (1975).

Layton and Sharifi (1978) examined the basic semantic structure underlying sentences produced by Down's Syndrome children and nonretarded children. Down's Syndrome children used the same semantic features as did the nonretarded children, but with much less frequency or consistency. For example, both groups

produced the four basic semantic verb categories (state, process, action, and process-action), but the retarded group used the process (The road ends) and the process-action (John cut the paper) verb forms less frequently than the nonretarded children did. In contrast, there was a tendency for the retarded children to use more state (The road is rough) and action (John runs) verbs than the normal group did. This finding was interpreted as indicating that Down's Syndrome children are more aware of or possibly restricted to the "here-and-now" aspects of their environment: "As children's cognitive abilities mature, their ability to perceive past and future events improve; consequently, their expressions will include more semantic concepts depicting these changes" (p. 444). Overall, Layton and Sharifi concluded that there are more similarities than differences in the usage of verbs by Down's Syndrome and normal children, and the primary cause of delayed semantic development among Down's children is their immature level of cognitive development.

Recent research on how children use words to convey meaning suggests that retarded and nonretarded children are more alike than different in this dimension. Young normal children have been shown to use single-word utterances for more complex purposes than the simple labeling of an object. For example, Greenfield and Smith (1976) suggest that young children call attention to a changing element or aspect which is undergoing the greatest amount of change. Leonard, Cole, and Steckol (1979) reported that retarded preschool children label objects in much the same way as the nonretarded children in Greenfield and Smith's study to convey the greatest amount of information efficiently.

Similarly, there is also some evidence to suggest that retarded children seem to impose meaning on verbal information in much the same way as nonretarded children do. Nonretarded children frequently rely on contextual information and their knowledge of real world relationships to interpret the meaning conveyed by a speaker. Dewart (1979) designed a study to determine whether mentally retarded children used strategies similar to those used by nonretarded children for interpreting the meaning of sentences. Three types of sentences were used that varied according to the kind of semantic cues they provided the child for use in interpretation: neutral, probable, or improbable. Sentences that were neutral with respect to semantic expectations would make sense if either noun served as the subject in the sentence. Probable sentences described likely events, while improbable sentences described events that were unlikely to occur in the real world (e.g., The gate jumped the horse). If children were using semantic constraints to interpret these sentences, they would respond correctly more often to probable sentences than to neutral sentences, and least often to improbable sentences. Dewart concluded that retarded children are able to use contextual information and previous knowledge concerning the likelihood of a given event to guide their understanding of sentences in much the same way as nonretarded children of a comparable mental age do.

In summary, mentally retarded persons appear to demonstrate less mature lexical development than normal children do. They show less abstraction, fewer other-references, and greater attention to semantic features that are anchored in the present. They typically define words in more simplistic and concrete terms because abstraction develops much later in retarded populations due to cognitive

limitations. Some studies indicate that retarded persons have a smaller lexical store than normal. Even when they are shown to be superior in vocabulary size, they typically demonstrate less variety and diversity in their lexicon. Despite these differences, the sequence of acquisition for lexical items has been shown to be the same as in normal development, only slower. The differences tend to diminish as mental age increases in the retarded population.

In general, mentally retarded persons appear to be able to understand and use semantic information in much the same way as nonretarded persons of the same mental age do. They have been shown to use the same types of semantic features in their speech as nonretarded children, even though they use them less frequently and with less consistency. The retarded also tend to use similar strategies for interpreting meaning and conveying information to others. Therefore, the semantic system of mentally retarded children develops normally, although they tend to lag behind in their rate of acquisition. Retarded persons can generally be expected to perform much like younger normal children do, sometimes even more immaturely than their mental ages would cause one to predict. Consequently, in developing interview questions for mentally retarded consumers with low mental ages, one should, to the extent possible, use very simple sentence constructions which draw on concrete and commonly used words.

LANGUAGE STRUCTURE OF THE MENTALLY RETARDED

The structure of language used by mentally retarded individuals has received increasing attention in the literature over the past several years. This aspect of language acqui-

sition includes the areas of morphology, or rules for forming words, and syntax, or rules for forming grammatically correct and meaningful sentences. Recently, there has been increasing emphasis on the mentally handicapped person's ability to comprehend language structures apart from his or her expressive abilities. In any attempt to obtain information from mentally retarded consumers about their needs, an understanding of their limitations in comprehending speech would be of great importance, since expressive requirements could be minimized by using various structured interviewing strategies.

Morphology

Morphology is the study of word forms in a language. Berko's (1958) test of English morphology has become the most widely used method for studying the development of morphological rules in children, and has also been extended to investigating the process of morphological rule-learning in the retarded. For example, the child is shown a picture of an imaginary object and is told "Here is a mook; now there is another one. There are two of them. There are two ____." The child must provide the correct plural form of the nonsense word.

In reviewing the studies which employed Berko's technique with retarded children, Yoder and Miller (1972) and Cromer (1974) have concluded that retarded subjects matched on mental age learn the rules of morphology in a similar manner to normal children, but at a slower pace. However, some researchers contend that the retarded also differ from normals in their ability to generalize these rules to new, unfamiliar stimuli (Leiber & Spitz, 1976; Lovell & Bradbury, 1967).

A direct test of the rule-learning deficit hypothesis was made by Bartel

(1970) with a group of mildly and moderately retarded children. She predicted that if such a deficit existed in the retarded population, moderately retarded children would improve on specific words when trained to provide the correct inflections for them, but the improvement would not generalize to other words. Contrary to her hypothesis, the two IQ groups in the study generalized to untaught items equally well.

In light of evidence provided by Dever (1972) on the predictive validity of Berko's procedure with the mentally retarded, much of the preceding information must be interpreted with caution. Consistent with previous findings, he and Gardner (1970) found that retarded children did not perform as well as normals on the Berko test even when matched on mental age. However, they observed that during spontaneous conversation some of their retarded subjects demonstrated appropriate use of many morphemes, even though they had performed poorly on the Berko task. Dever (1972), therefore, used a revised version of Berko's test to determine whether performance on either part of Berko's test could predict the free speech errors of retarded subjects. Since it could not, Dever concluded that the Berko test is not a valid diagnostic test of morphology usage in the mentally retarded.

However, it is important to note that retarded and normal groups found the same items to be relatively difficult in Dever and Gardner's study. Plurals and possessives were easiest, followed by progressive verb forms. Irregular verb forms were not handled well by any of the children before mental age of 10. Therefore, despite the lack of predictive validity of this test with respect to morphological usage in spontaneous

speech, it appears that retarded children develop understandings of morphology in the same sequence normal children do. Overall then, while the mentally retarded may sometimes demonstrate less flexibility in generalizing morphological rules to new situations than normal children, there seems to be no strong support for the notion of a rule-learning deficit in retarded children. In addition, Graham and Graham (1971) reported that even severely retarded persons produce a considerable amount of speech which is formulated on the basis of appropriate grammatical rules and operate at a level comparable to that of normal children, of the same mental age.

Syntax

In studies of syntax development in children one method that is frequently used is the cloze procedure, which requires the child to supply a word missing from a sentence. Performance on this task is determined by the child's ability to correctly judge the grammatical class of the word and then to select an appropriate word from that class to complete the sentence. Research has demonstrated that retarded children are more likely to use sequential strategies rather than paradigmatic strategies in making word associations (Simmel, Barritt, Bennet, & Perfecti, 1968). That is, like young normal children, they associate words on the basis of contiguity (e. g., red-blue). For this reason, Simmel, Barritt, and Bennet (1970) predicted that retarded subjects would find the cloze procedure more difficult than the nonretarded children. They also predicted that when a word was omitted from the end of a sentence, retarded children would succeed more often since more semantic clues would be provided to guide the selection of the missing word.

Retarded children ages 10 to 14 years with an IQ range of 60 to 80 were matched with two groups of nonretarded children: one on chronological age and one on mental age. In general, the retarded children did more poorly than either nonretarded group in the use of specific grammatical form classes. As predicted, the retarded children also showed greater improvement in performance when supplying words omitted from the end of a sentence than normal children, suggesting that the retarded children were more dependent on sequential strategies for selecting words. However, in reviewing this study, Cromer (1974) observed that there was sufficient similarity in the performance of the retarded children and the normal children matched on mental age to suggest that the retarded group was not responding with deviant strategies. For example, both the retarded and nonretarded children found nouns the easiest to replace.

Goodstein (1970) replicated Semmel et al. (1970) study with the addition of a recognition version of the cloze procedure. These results indicated only quantitative differences between retarded and normal children. Both studies, then, suggest that sentence complexity affects sentence comprehension among the retarded in much the same way as it does among normal children. Retarded children did not make different kinds of errors, just more of them. Consistent with the findings of Semmel et al. (1970) Goodstein found that the order of sentence difficulty was the same for both the retarded and nonretarded children.

Several researchers have attempted to determine if the kinds of sentence constructions that normal children and adults have difficulty understanding also cause problems for the retarded. Semmel and Dolley (1971) studied a sample of Down's Syndrome children ranging in age from 6 to 14

years with an IQ range of 22 to 62. They presented their subjects with both comprehension and imitation tasks using the four sentence types, and expected to replicate the order of sentence difficulty reported by Slobin (1966) for normal children and adults. In the comprehension task, subjects were presented with a pair of pictures, each representing one aspect of a reversible situation (e.g., A clown hitting a ball and a ball hitting a clown). As predicted, Down's Syndrome children performed best on the simple declarative sentences (slightly above chance). Performance on passive and negative passive sentences was only at a chance level of responding. Surprisingly, Down's Syndrome children performed below chance on negative sentences. The authors suggested that these children were responding to negative sentences as if they were simple affirmative sentences, and proposed that Down's Syndrome children may be unable to process negative sentences, or may simply ignore the negative marker in these sentences. As a result, Slobin's findings were not replicated in this group of retarded children.

In contrast, Lamberts and Weener (1976) reported that over half of the retarded subjects in a language-training project successfully repeated sentences which contained a negative marker. Since previous research has demonstrated that retarded children correctly imitate only those sentence structures which already exist in their own spontaneous speech patterns or which they are able to understand (Lackner, 1976; Lenneberg, 1967), the hypothesis that retarded children cannot process negative sentences seems less plausible. Moreover, in the Lamberts and Weener study, comprehension of negative and affirmative sentences was related to sentence reversibility (that is, to whether subject and object can be interchanged, as in "the

cat chased the dog," or whether they are irreversible as in "the cat climbed the tree"). Nonreversible negative sentences were actually easier to understand than reversible affirmative sentences for retarded persons with a mental age of five or above.

Semmel and Dolley (1971) did not find sentence comprehension to be related to level of intelligence. However, later research has reported that the ability to comprehend sentences in retarded children does vary with mental age (Bartel, Bryen, & Keehan, 1973; Lamberts & Weener, 1976; Walker, Roodin, & Lamb, 1957). Lamberts and Weener note that, on the average, the subjects in Semmel and Dolley's sample were much younger than their sample of retarded children. Therefore, comprehension of more complex sentence constructions may not have developed in Semmel and Dolley's Down's Syndrome children because a requisite level of cognitive or linguistic functioning had not been attained. Semmel and Dolley (1971) did find level of intelligence among Down's children to be strongly associated with performance on a sentence imitation task.

Mittler (1970) proposed that sentence comprehension depends more on structural complexity than on sentence length. He reported that the rank order of sentence difficulty was similar in a group of retarded children and a matched sample of nonretarded children. An analysis of errors suggested that comprehension was most affected by the structural complexity of the sentence in both groups of children. Berry (1972) examined the effects of structural complexity on sentence comprehension when sentence length was controlled. He compared the comprehension ability for retarded and nonretarded children for simple declaratives (e.g., The boy is fishing here.) and more com-

plex possessive sentence constructions (e.g., The boy's fish is here.) Possessive sentences were more difficult for both groups of children.

A number of other studies have shown that retarded children are generally able to process simple declarative sentences with the same accuracy as nonretarded children of a comparable mental age. However, as transformational complexity increases, the retarded child lags further behind his or her mental age counterpart in sentence comprehension. For example, Dewart (1979) reported that retarded children above mental age three comprehend simple active sentences above chance, while passives are interpreted at a level below chance. She indicated that their performance was very similar to that of a group of nonretarded children whose average mental age was lower, which supports the idea of delayed acquisition of passive constructions in the retarded. Retarded children with a lower mental age, however, performed at a chance level on both active and passive sentences. Therefore, Dewart concluded that, prior to mental age three, retarded children are quite limited in their ability to understand and use syntax.

Despite the fact that the acquisition of more complex sentence structures is delayed in the language development of the mentally retarded, such learning does occur. The order of acquisition of complex sentence constructions parallels that observed in the developmental sequence of normal children (Berry & Foxen, 1975; Gallagher, 1969). For example, Graham and Graham (1971) collected language samples of nine institutionalized retarded males, ages 10 to 18 years, with mental ages ranging from 3-6 to 10-0. Retarded males with lower mental ages produced most of their sentences through the use of

base strings or kernel sentences with no transformational rules applied. Mental age was highly correlated with use of transformations to form more complex sentences. Thus Graham and Graham concluded that language facility does depend primarily on intellectual level in retarded persons and that retarded children develop the rules of language in much the same manner as nonretarded children only at a slower rate.

In one of the most linguistically sophisticated studies to date on the language development of retarded children, Lackner (1976) has provided additional support for this viewpoint. He collected language samples from five retarded children who represented five ascending levels of intellectual functioning and compared their linguistic capabilities with those of nonretarded children. The order of appearance of sentences in the language patterns of the retarded children was consistent across the five mental age levels and followed a regular order of increasing structural complexity: declaratives, negatives, questions, negative questions, passives, negative passives, and negative passive questions. No sentence type of a given order of complexity was found in the retarded child's grammar unless all lower order constructions were also part of his or her language repertoire. More significant was the finding that as mental age increased the number of sentence transformation roles understood and used by each child also increased, indicating that language development as measured by transformational complexity does improve with higher intellectual functioning. This supports the previous findings of Graham and Graham (1971). Lackner also observed that none of the sentences of the retard-

ed children were incompatible with normal adult usage, supporting the idea that retarded persons do not use deviant language structures. Retarded and normal children exhibited similar developmental trends, but the language development of the most severely retarded children appeared to be arrested at a level lower than that of their mental age-matched counterparts.

There has been much debate in the literature over the relationship between comprehension and production in the language acquisition process. Current thinking assumes that comprehension precedes expression or production in the developmental sequence. Research on language development of nonretarded children supports this notion (Fraser, Bellugi, & Brown, 1963; McCarthy, 1954). Other recent studies with nonretarded children have also demonstrated that prior training in comprehension skills facilitates acquisition of production abilities (Vasta & Teitelbaum, 1976; Whitehurst, 1977).

In contrast, some studies with mentally retarded persons have reported different and conflicting information. For example, Guess (1969) and Guess and Baer (1973) concluded that language reception and production in mentally retarded children may be independent processes. They found that training retarded subjects in one modality did not produce generalization effects in the other language modality. Similarly, Miller, Cuvo, and Borakove (1977) found that teaching reception skills to the retarded did not improve expression; however, training in production did improve receptive skills. This research suggests that a qualitative difference may exist between retarded and nonretarded populations with respect to the direction of generalization between comprehension and production.

Lackner (1976), however, feels that production of grammar lags behind comprehension in the mentally retarded. He observed that both retarded and nonretarded children in his study frequently modified complex sentences on an imitation task in such a way that the meaning of the sentence was retained but sentence complexity was reduced. For example, a child would sometimes repeat the passive sentence construction "Mary was hit by the ball" as "The ball hit Mary." Lackner felt that this indicated that the child's productive skills for grammar were not as well developed as his ability to understand syntax.

Similarly, Walker, Roodin and Lamb (1975) reported that expressive language consistently lagged behind comprehension in a sample of retarded children with a mental age of 3 to 7.5. In a group of nonretarded children with a comparable mental age range, this difference occurred for only the youngest children, ages 3 to 3.5. As mental age increased in the retarded group, the magnitude of the difference was found to decrease so that by mental age seven expressive and receptive abilities were nearly equal. The authors concluded that expressive tasks may place a greater demand on short-term memory than comprehension tasks do. In their research, "maximal estimates of language ability in retarded subjects were obtained in tests that measure comprehension without requiring verbal production by the child" (p.551). Consequently, one might expect that maximal information could be obtained from retarded persons using interview questions which make minimal demands for verbal production.

In summary, then, the research evidence suggests that the language

structures of mentally retarded persons are quite similar to those of nonretarded children at a similar mental age. In addition, there is no strong evidence to suggest that the mentally retarded use language structures or strategies that are incompatible with those of normal children, even though language acquisition frequently occurs at a much slower rate. Language delays are likely to be greatest for retarded persons below mental age seven, since prior to this level of functioning they do not seem to possess the cognitive capabilities necessary for comprehending the more complex aspects of grammar. In addition, retarded persons below mental age three would be expected to have very little understanding of correct grammatical usage. Since retarded persons have been shown to consistently perform best on language tasks involving simple sentence structure, this would suggest that unnecessary complexity should be avoided when questions are asked of the retarded.

Question Development in Children and Its Implications for the Retarded

No research that we are aware of has directly examined the development of question-answering competence in retarded individuals. However, research on normal children's comprehension and use of questions has been reported within the last 15 years, and it is highly relevant to our concern with how to phrase questions. The evidence suggests that there is a basic sequence of question development in normal children, although individual studies have shown some variation in the order of mastery.

It seems reasonable to expect that a similar pattern of question development occurs in the mentally retarded, since their language processes have not been shown to be qualitatively different from those of nonretarded children.

By age four, children are typically able to respond appropriately to most question forms. Bellugi (1965) observed that before age two children do not seem to understand many of the wh-questions. Comprehension for these question formats begins to emerge somewhere around 18 to 28 months of age, with yes-no, what- and where-questions generally being the first to be understood. Why- and who-questions were among the next questions to be acquired (around the age of 32 months). These observations correspond very closely with an early study by Smith (1933), who found that what, where, and yes-no question types were the most frequently produced by children ages 1-6 to 6-0.

Ervin-Tripp (1970) conducted the first major study of how children comprehend questions. She observed the responses to questions of children ages 2.5 to 3.1 over a one-year period. Initially, all children could answer what, where, and yes-no questions appropriately, which agrees with Bellugi's observations. The relative order of acquisition for other wh-questions, according to Ervin-Tripp was: why and who-subject; how and where-from; when and who-subject questions. However, since this order was not stable across all children in the study, she presents this as a tentative developmental sequence for question acquisition in normal children.

The level of the child's cognitive development appears to influence

what questions he or she can answer. For example, around the age of three, most children begin to produce adequate causal responses to simple why questions that are manageable at their level of cognitive functioning (Bellugi, 1965; Ervin-Tripp, 1970). By comparison, the child's understanding of time concepts does not develop fully until after age six (Piaget, 1927), which would explain why when-questions are among the most difficult for young children to understand.

Recently, Tyack and Ingram (1977) attempted to replicate the findings of Ervin-Tripp in a study of children's question comprehension and production. The children in the production experiment used yes-no, what, and where questions most frequently, with the exception of 3-6 to 3-11 group who used why and how questions more frequently than where questions. The authors suggested that what and where questions are used frequently by young children because they are closely tied to the child's immediate experience. An alternate explanation has been suggested by Hooper (1971), who proposes that yes-no and labeling questions like "What is that?" are much easier for children to answer than open-ended and explanatory questions such as how and why because they make fewer demands for grammatical processing and verbal elaboration. Possibly both this linguistic explanation and the cognitive development viewpoint account for the earlier use and comprehension of yes-no and what questions in children.

While frequency of use does not necessarily reveal the order of question acquisition in children, the usage figures reported by Tyack and Ingram (1977) do agree with the

comprehension data of Bellugi (1965) and Ervin-Tripp (1970). However, the acquisition sequence for question comprehension which emerged from Tyack and Ingram's sample of 3.0 to 5.5 year olds was somewhat different from that reported in earlier studies. The order of question type was as follows: where-intransitive (94%); why-intransitive (83%); why-transitive (81%); who-subject (80%); where-transitive (67%); what-object (57%); who-object (56%); when-intransitive (52%); and when-transitive (48%); how-transitive (38%); how-intransitive (35%), and what-subject (35%). According to these findings, intransitive verb forms are better understood than transitive forms, except in how-questions. As the authors note, when-questions were easier to answer than how-questions for these subjects, where this was not the case in Ervin-Tripp's study. A more obvious discrepancy was that what-questions were more difficult to answer in this study. In fact, what-subject questions were actually the most difficult to understand for these children. The authors suggest that young children may automatically associate what questions with the object. More generally, Tyack and Ingram concluded that transitivity and the semantic features of a verb are as important in determining the child's response to a question as is the question format used.

The research evidence on the acquisition and comprehension of questions in normal children seems to suggest that the mentally retarded, especially those who are severely retarded, would find what, where, and yes-no questions the easiest to comprehend and answer. They would probably find who, how, and why questions more difficult to comprehend and

even more difficult to answer because of the greater demands for verbal expression made by these question formats. When-questions would probably be the most difficult to answer because of the higher level of cognitive functioning necessary to understand time concepts. However, as Tyack and Ingram have indicated, the transitivity and semantic features of the verb used could also affect this order of difficulty significantly. Since level of cognitive development appears to be an important factor in the normal child's ability to respond to questions, one would expect retarded children to acquire these question forms at a slower rate, especially if they are severely retarded.

Response Effects in Survey Research

In any research, including research using interview techniques, validity of measurement is a critical issue. Simply stated, one needs to know what an interview measures; that is, whether responses to an interview accurately portray the actual behaviors or attitudes of the individual interviewed. Since part of the purpose of our research was assessment of the validity of information obtained from mentally retarded respondents, we sought guidance from the literature on measurement and measurement error in survey research.

Traditionally it has been assumed that responses to an interview are composed of two factors. The first is the respondent's "true answer," or the actual information sought by the interviewer. The second is response error, or variability in the respondent's answer due to all sources other than the "true answer." Response error, then, is the difference between the "true

answer" and the answer actually given in the interview. Response error arises from many sources; a response effect is the response error due to one specific factor. If response effects tend to cancel one another out so that mean response error is zero, straightforward interpretation of the interview data for a group is possible. If, however, some response effects are large enough that the obtained answer consistently varies in a given direction from the "true answer," interpretation of the interview data is much more difficult. Data from our research suggest that such powerful response effects may be at work in interviews with mentally retarded individuals, tending to reduce the validity of answers. The literature review which follows is an effort to assess the relative importance of response effects due to various factors, with particular emphasis on those variables expected to have a special importance in interviews with mentally retarded persons. We will be concerned with learning whatever lessons may be learned from survey research with the general population that might aid in interviewing mentally retarded respondents.

Fortunately, Sudman and Bradburn (1974) have provided a review and analysis of research of response effects which is a landmark in the field. These authors used a coding strategy to combine results of an enormous number of interview studies and estimate the relative importance of various response effects. The present review is, in many respects, a summary of the major conclusions of the Sudman and Bradburn work, with emphasis and elaboration of those biases that seem potentially most relevant to

interviewing mentally retarded persons. In particular, the model of the interview situation and the conceptual framework for evaluating different types of response effects are adopted directly from Sudman and Bradburn. This model views the interview as composed of three elements: the task of giving and receiving information, the role of the interviewer, and the role of respondents. Thus, the model regards response effects as a function of either task variables, interviewer variables, or respondent variables.

Task variables are the form and conditions of the interview itself. The specific characteristics of the task may be expected to exert a tremendous influence on the nature of the information obtained. Question wording is an important task variable. Bias may be introduced directly by the wording of a question, or may result indirectly from the length or difficulty of questions. One might also expect the structure of questions (e.g., closed-ended vs. open-ended) to exert an effect on responding. Questions which offer the possibility of an agreeing response can give rise to powerful acquiescence effects, a tendency to agree or say "yes" regardless of question content. Another task variable is the degree to which an interview raises issues of self-presentation or, more specifically, creates in respondents the desire to present themselves in a favorable light. Whether an interview is directed towards behavioral or attitudinal information is another task variable of interest. Still another may be described as the informational demands of the interview. Questions may place greater or lesser demands on the respondent's

memory, and may pertain to topics of greater or lesser salience to the respondents. Finally, the method of administration (for example, face-to-face vs. self-administered interviews) is potentially important.

Interviewer and respondent variables generally refer to characteristics of the individuals involved, such as their age, race, sex, and level of education. The magnitude of response effects due to interviewer characteristics is usually dependent on respondent characteristics and the topic of the interview. Similarly, the importance of respondent variables typically depends on interactions with task and interviewer variables. In fact, although for the sake of clarity Sudman and Bradburn describe task, interviewer, and respondent variables as separate sources of response error, their data show clearly that most important response error is a function of combinations of these variables. For example, while question wording and the education of the respondent can each give rise to response effects, the response effect due to the interaction of these two variables is larger.

TASK VARIABLES

The trend in survey research has long been to neglect task variables in favor of investigations of interviewer and respondent characteristics. The most general conclusion of the Sudman and Bradburn review is that this emphasis is unwarranted; in fact, the most general and important response effects arise from task variables. The specific task variables found to be of importance are evaluated below. Much of our own research effort focused on the effects of task variables such as question structure and wording.

Question Wording

The specific wording of a question might be expected to give rise to response effects in two general ways. First, bias inherent in the wording, as in the use of emotionally laden or leading words, may alter the respondent's answer. Second, difficult or ambiguous questions may enhance response effects due to other variables. Essentially, if the respondent cannot interpret the question, he or she cannot give the "true answer."

Good evidence exists that responses can be systematically biased by question wording. For instance, Rugg (1941) asked the questions: "Do you think the United States should allow public speeches against democracy?" and "Do you think the United States should forbid public speeches against democracy?" In response to the first question, 62% of respondents believed such speeches should not be allowed, while in answering the second, only 46% were willing to forbid speeches. Clearly, the wording of these questions had a strong influence on the response effects due to question wording.

Schuman and Presser (1977) replicated and extended these results. Using Rugg's original questions, they found response effects due to questions wording to be of similar magnitude and direction. Additionally Schuman and Presser challenged the assumption, implicit in much response error literature, that question wording and respondent characteristics do not interact. If this assumption were correct, then response effects due to question wording would not affect comparisons of subgroups of respondents. However, Schuman and Presser showed that less educated respondents were

more influenced by question wording than were more educated respondents. Among respondents with 13 or more years of education, the "allow" and "forbid" forms of the question yielded estimates of the proportion of respondents opposing speeches which differed by only 5.9%. This difference was 25.6% for respondents with 0-11 years of education. In short, question wording can have powerful response effects, and these response effects are greater for less educated respondents. These results may have important implications for the interviews with mentally retarded individuals, who are at the low end of the educational continuum.

The length and difficulty of questions also may affect responding. Length of question, simply defined as the number of words, has generally had no overall effect on responding in face-to-face administration. Of more interest for our research is the interaction of question length with respondent variables. Response effects for respondents with a high school education or less are largest for questions more than 18 words long. Elementary students are particularly influenced by long questions. This finding may have important implications for interviewing mentally retarded persons; if long questions cause large response effects in younger and less educated respondents, it may be particularly important that questions posed of mentally retarded individuals be short and clear.

The number of letters per word has been used as an indicator of question difficulty. Difficulty of questions has no general effect for attitudinal items; although response effects for threatening behavioral items increase for questions using longer words.

Question Structure

The issue of question structure is typically posed as a choice between closed-ended questions, which ask the respondent to choose from a fixed set of response alternatives, and open-ended questions. Although Sudman and Bradburn found no consistent effect due to question structure, they did find that response effects due to issues of self-presentation are larger for closed-ended than for open-ended questions. That is, for threatening topics or topics for which there is a socially desirable answer, closed-ended questions increase response effects.

In addition, direct comparisons of closed and open-ended surveys have demonstrated that different forms of a survey can yield very different results. Jenkins (1935) reported a study which involved construction of an exhaustive checklist by using all responses to a previously administered open-ended survey. This checklist (including an "all others" items) yielded response patterns closely comparable to the open-ended survey. However, when some of the popular items were deleted from the checklist, response patterns to the incomplete checklist form differed radically from responses to the open-ended form. It is at least clear that the make-up of a checklist can have important influences on responding, and that if items often mentioned in response to open-ended questions are omitted, under-reporting may result.

Belson and Duncan (1962) compared checklist and open-ended surveys with regard to TV programs watched and periodicals read by respondents during a specified period preceding the interview. They reported far higher claims of activity in response

to the checklist form than to the open-ended survey, with the size of differences between question forms varying across items. Also, Belson and Duncan included some checklist items naming TV programs not actually televised during the period of inquiry. A small fraction of respondents claimed to have watched these programs, demonstrating that in at least some instances closed-ended surveys can cause overreporting.

Schuman and Presser (1977) also reported large differences on some items between endorsement rates on open and closed forms. Response effects due to question structure were found to be related to the education of the respondent. For example, the poorly educated were more likely than other respondents to endorse a job security item on the closed-ended form, but less likely to list the same item on the open-ended form. Response effects were also larger for the less educated in a comparison of agree-disagree vs. forced choice questions.

In summary, although Sudman and Bradburn observed no general response effect due to question structure, direct comparisons reveal substantial response differences between closed and open-ended questions. Closed-ended (checklist) questions tend to provide a higher yield of claims, which in some instances may constitute overreporting. Finally, response effects due to question structure may be exaggerated in less educated respondents, another hint that special problems may arise in interviewing retarded persons.

Acquiescence

Questions which allow for the possibility of an agreement response

"yes," "true," "agree," and the like, may give rise to acquiescence, or a general tendency to agree independently of item content. Since acquiescence was examined in the present study, we will consider it in some detail.

Unlike most response effects discussed thus far, acquiescence is a topic of interest primarily within personality research. The large body of literature on acquiescence (see Cronbach, 1946, for his influential introduction of the issue) emphasizes the effects of acquiescence on the interpretation of personality inventories like the California F scale and the MMPI. For example, items on the F scale are all keyed true; it has been suggested (see, for example, Bass, 1955) that the F scale identified particularly acquiescent individuals, rather than authoritarian individuals as argued by the creators of the scale. However, this body of literature arrives at no consensual conclusion regarding the importance of acquiescence in personality inventories (see Rorer, 1965, for a critical review of the literature).

No consensus has been reached on this issue largely because it is a methodologically difficult one. In order to evaluate the importance of acquiescence in any instrument, one must somehow separate question form and question content--ordinarily highly confounded factors. This separation is particularly difficult with personality inventories, where the "true" answer to an item is now known and where items are usually attitudinal and rather complex. It is therefore not particularly surprising that the role of acquiescence in these inventories is unclear.

Our research is directly concerned with acquiescence as a factor which might invalidate responses to questions much less ambiguous than those

typically found on personality inventories. The questions we are concerned with tend to be behavioral (Do you date?) or relatively simple attitudinal (Are you usually happy?) questions. As we will demonstrate later, acquiescence may be a highly significant biasing factor in interviewing mentally retarded persons. The research described below is organized around four approaches to assessing acquiescence.

Content-free Measures

One strategy for demonstrating the effect of acquiescence is to employ "interviews" which are essentially content-free. In this situation, any non-random response patterns observed must be responses to the form rather than the content of the question. For example, Berg and Rapaport (1954) gave subjects a set of response alternatives without any questions. They were instructed to try to guess the correct answers. Berg and Rapaport observed clearly non-random patterns of responding to this essentially content-free measure. Of particular interest is the fact that response alternatives "yes," "satisfied," "true," and "agree" were significantly favored by respondents regardless of the order in which alternatives were presented.

In a similar study, Gerjouy and Winters (1966) tested response preferences in 60 institutionalized educably retarded adults. They presented the subjects with pairs of identical geometrical figures and asked if the figure to which they pointed was the larger (or smaller). Of all responses, 59.1% were "yes," significantly different from chance at the .001 level.

Clearly, a preference for agreement responses exists in both normal and retarded individuals in an extremely unstructured situation. It is conceivable that questionnaires may essentially be unstructured or "content-free" for mentally retarded respondents if the demands of the questions exceed intellectual abilities; in such a case, responding would probably tend to resemble the acquiescence observed in the above studies.

Tests with Known Answers

A second strategy, one employed in the present study, is to use items to which the correct answers are known. If in this situation the pattern of responding differs systematically from the pattern of correct answers, a response set may be inferred. Cronbach (1942) used this approach with true-false tests and observed that students in general are inclined to guess "true" when in doubt. Further, Cronbach found that each student's tendency to acquiesce was somewhat consistent across tests; a given student's tendency to answer "true" correlated from .36 to .61 across tests. Interestingly, if students guess "true" when in doubt, then false items have greater validity than true items as measures of the students' knowledge. In fact, Cronbach showed that in some cases the false items have greater validity than the true and false items combined. This finding may be important in survey research with particularly acquiescent populations; from such persons, a "no" response would, on the average, be more meaningful than a "yes" response. In short, individuals have a somewhat stable tendency to answer "true" when in

doubt, and this fact tends to make false items more valid than true items.

The Item-reversal Approach

Unfortunately, neither of the above research strategies can be applied in most practical contexts. In interview situations, items have content and the correct answer is almost always unknown. The item-reversal approach has been devised for use in such circumstances. If an item can be administered in two forms so that the content is precisely reversed in the two administrations, self-contradictions in which the respondent agrees with both statements are a clear indication of acquiescence. For example, if a respondent replies "yes" to both "I earn more than \$5000 per year," and "I earn \$5000 or less per year," response is clearly based on question form rather than on question content. This technique gives the most definitive demonstrations of acquiescence, and was employed in the present study. Its only major difficulty is that clear content reversal of many items, especially complex attitudinal items, is virtually impossible. Rorer's (1965) critique of item reversals with the F scale makes this point quite convincingly.

Hare (1960) administered original and reversed forms of a 44-item interview with an "agree"-"disagree"-"can't answer" response format. Subjects were black and white women. Self-contradiction was observed, particularly among black women, who gave contradictory answers to an average of 20 of the 44 pairs of items. Hare did not separately report rates for contradictions of the agree-agree and disagree-disagree types (a

common flaw in item-reversal research), but did report that contradictions were "mostly" of the agree-agree type. This result indicated that a tendency on the part of certain types of respondents to acquiesce can generally invalidate interview data.

One study directly investigated the effects of acquiescence with mentally retarded respondents. Gozali and Bialer (1968), working with a personality inventory called Children's Locus of Control Scale, reversed the 23 items on this scale and administered both forms to a sample of 189 mentally retarded individuals of both sexes. Subjects ranged in age from 16 to 30 years and in IQ from 58 to 91. The measure of locus of control derived from the standard form of this test was highly correlated with the measure of locus of control derived from the reversed form; that is, responding to the scale appeared to be based on item content rather than item form. Unfortunately, Gozali and Bialer reported only correlation coefficients, mixed questions keyed "yes" and "no" on both forms, and did not report self-contradictions on an item-by-item basis, leaving us uncertain about the extent to which acquiescence might have been operating.

Lenski and Leggett (1960) imbedded a single question and its reversal. ("It's hardly fair to bring children into the world, the way things look for the future," vs. "Children born today have a wonderful future to look forward to,") in a long interview, separated by about one-half hour. For the sample as a whole, 8% of respondents agreed with both items. This tendency to acquiesce was related to both education and race. Overall, 20%

of black versus 5% of white respondents agreed with both statements. For white respondents, only 2% of those with some college education acquiesced, while 9% of those with 8th grade education or less acquiesced. Among black respondents, 14% of those with some college agreed with both items, compared to 32% of those with 8th grade education or less. Thus, black respondents and less-educated respondents showed a more marked tendency to acquiesce.

Rothenberg (1969) used item-reversal in studies of cognitive development in young children. Studying conservation of number, Rothenberg asked his subjects both "Does this bunch have the same number of blocks as this bunch?" and "Does this bunch have more blocks?" Subjects were 4- and 5-year-olds, divided into lower and middle class (distinguished by neighborhood). This class distinction also divided the sample according to race, for the lower class group consisted of black and Puerto Rican children, while the middle class sample included no Puerto Ricans and only 4% blacks. Additionally, the class distinction appears to have divided the sample according to intelligence. As measured by the Peabody Picture Vocabulary Test, the 4- and 5-year-old lower class children had average IQ's of 72.2 and 72.5, respectively. The middle class 4- and 5-year-olds had IQ's of 110.6 and 101.9. On a single test of conservation, 65.1% of lower class 4-year-olds contradicted themselves by answering "yes" to both forms of the question stated above. The acquiescence rate was 40.0% for lower class 5-year olds, 7% for middle class 4-year olds, and 10.0% for middle class 5-year olds. The difference between average acquies-

cence for the lower middle class samples was significant at the .001 level. Because race, social status, and IQ are confounded in this design, it is impossible to determine which of these variables was responsible for the difference in acquiescence rates in the two groups. However, this is a clear demonstration of the potentially ruinous effects of acquiescence on validity. For the sample as a whole ($N = 210$), 27.8% of all responses were yes-yes self-contradictions. (This far outweighs the incidence of no-no self-contradictions, at 5.2%.) The fact that this very high acquiescence rate occurred among young respondents and was especially pronounced for respondents of low intelligence at least suggests that acquiescence may be a particular problem in interviewing mentally retarded individuals.

Measures of Individual Differences in Acquiescence

A fourth approach to the evaluation of acquiescence involves the construction of scales specifically intended to measure the acquiescence tendencies of different individuals. The development of such scales typically involves constructing a group of items which is held to be heterogeneous with respect to content. Furthermore, items are balanced so that equal numbers of agreeing and disagreeing responses are required for expression of an extreme preference within any content area. It is assumed that the influence of content on the total number of agreement responses is thus eliminated, and the agreement score is taken to be a measure of inclination to acquiesce. In essence, this is simply an imprecise form of the item-reversal strategy. Couch and Keniston (1960) created such an acquiescence scale, con-

sisting of 360 items drawn from a variety of personality instruments and using a Likert type scale. Couch and Keniston attempted to correlate acquiescence as a personality variable with a number of other variables, constructing a personality profile for the yeasaying respondent. For our purposes, their most significant finding was that yeasaying and intelligence are unrelated. However, Couch and Keniston used crude measures of intelligence (the intelligent-dull scale of the 16 P.F., and the College Entrance Examination Board scores) and used a sample with a narrow range of intelligence (subjects were college undergraduates). Thus, it is not surprising that yeasaying and intelligence were unrelated in this study.

Wells (1963) derived a 20-item scale purported to measure acquiescent tendencies based on the earlier scale constructed by Couch and Keniston (1960). Wells reported that yeasayers, as defined by response to his scale, generally tend to endorse a wide variety of behavioral items. Further, Wells found that yeasayers tend to over-report in surveys investigating such topics as possession of magazines and recall of advertisements. Finally, it is interesting to note that Wells found yeasaying particularly common among younger and less-educated respondents.

It is tempting to interpret this latter finding as lending strong support to the findings of the item-reversal literature. However, this individual differences research must be interpreted very carefully; the imprecision of the content balancing procedures opens these scales to other interpretations. Inspection of the items of the Wells

(1963) scale suggests that rather than being "content balanced," this scale may in fact test for an attitude of optimism or uncritical enthusiasm. The conclusions of the acquiescence scales literature cannot be readily generalized to the issue of acquiescence as a response set tending to invalidate interview data.

Summary and Conclusions

There is a striking body of evidence attesting to the fact that interview data can be invalidated by a tendency on the part of respondents to give agreement responses independently of the content of the question. The studies using content-free measures suggest a possible mechanism for this effect; when respondents lack more substantive bases for responding, they tend to prefer agreement responses. The data drawn from situations where the correct answer is known support this conclusion and suggest additionally that answers to false items are more valid than answers to true items. The most impressive data on acquiescence come from item-reversal studies. These studies demonstrate clearly that acquiescence can operate powerfully enough to seriously reduce the validity of interview responses. Also, item reversal studies suggest that black respondents, less-educated respondents, and less intelligent respondents may be particularly likely to acquiesce. These latter two subject variables are, of course, of direct relevance to our efforts to interview mentally retarded respondents. Finally, data from individual difference measures of acquiescence have some indirect relevance to the present study. In short, acquiescence is potentially

a very serious problem in interviews, and the data suggest that the problem may be accentuated in interviews with mentally retarded populations.

Issues of Self-Presentation

Many investigators have suggested that distortion of responses may occur when the respondent seeks to portray himself or herself in a particular way. Sudman and Bradburn (1974) refer to such response effects as arising from "issues of self presentation." Within the personality assessment literature, much attention has been given to the influence of social desirability on responding. That is, respondents may tend to endorse items perceived as socially desirable and to avoid endorsing items perceived as socially undesirable, without regard to the "true answer." Thus, response patterns would reflect the perceived social desirability of the items, and not the actual behaviors or attitudes of the respondents. Sudman and Bradburn (1974) also raise the possibility that items of an extremely personal nature might threaten the respondent and alter responses.

Edwards (1957) presented data suggesting the importance of social desirability as a source of response effects. Subjects' scores on personality scales frequently correlate very highly with the rated social desirability of scale items. Edwards also argued that the correlation between two measures frequently reflects the degree to which they have similar properties on the dimension of social desirability.

Nevertheless, Sudman and Bradburn conclude that response effects attributed to threat and social

desirability are generally small. The interaction of self-presentation variables with other variables may, however, be important. For example, respondents are more likely to present themselves favorably when interviewed face-to-face than when afforded the relative privacy of a self-administered questionnaire. Thorndike, Hagen, and Kemper (1952) compared 500 self-administered and 500 face-to-face administrations of an inventory on psychosomatic symptoms. The respondents to the self-administered form reported some 15% more psychosomatic symptoms than did respondents to the face-to-face interview. When Knudsen, Pope, and Irish (1967) compared women's responses to face-to-face and self-administered interviews regarding the permissibility of premarital sex, 20% of respondents to face-to-face interviews said it was all right to have premarital sex, as opposed to 31% of the respondents to the self-administered form. In general, the magnitude of the effect due to method of administration increases as the probability of a socially desirable answer becomes greater.

As has been mentioned already, question structure also interacts with self-presentation issues to produce response effects. Closed-ended questions appear to increase response effects due to threat or social desirability. Furthermore, some interaction occurs between social desirability and respondent and interviewer characteristics. Sudman and Bradburn (1974) found response effects to differ as a function of the race and sex of both interviewer and respondent, depending on the degree to which an interview activated concern with self-presentation.

In short, issues of self-presentation in themselves have small response effects, but interact with other

variables to produce important response effects. In particular, method of administration interacts with social desirability; face-to-face interviews involving a strong social desirability component result in considerably less admission of information placing the respondent in a negative light. This finding may be important in interviews with mentally retarded persons, since these interviews are of necessity face-to-face. Mentally retarded respondents may tend to report attitudes and behaviors which they believe will meet approval, especially if they perceive the interviewer as having power. Social desirability also interacts to some degree with question structure and with the race and sex of the interviewer and respondent.

Behavioral vs. Attitudinal Information

One can distinguish two types of information sought in interviews. The first type is behavioral information, sought through such questions as, "In what year were you born?" The distinguishing characteristic of such information is that it can in principle be verified by reference to outside sources. Attitudinal information, on the other hand, involves the subjective valuations of the individual (e.g., "Do you like your current job?"). Such questions have no external referent; that is, they cannot, even in principle, be verified.

There is some suggestion in the survey literature that attitudinal information is more subject to response effects than is behavioral information. However, because attitudinal information has no external standard of accuracy, the measurement of response effects for behavioral and attitudinal

information differs fundamentally. Thus, response effects for the two types of questions cannot be directly quantitatively compared, and no general conclusion as to the relative susceptibility of the two types of questions to response effects is easily drawn.

Availability of Information

The availability to the individual of the information sought may have powerful effects on response validity. That is, if information is not very salient or is for some reason difficult to recall, the respondent's answer is likely to be biased.

Sudman and Bradburn (1974) showed that response effects differ greatly according to the length of the recall period involved in reporting behavioral information. Respondents tend to overreport behaviors for short recall periods and to underreport as length of recall increases. Length of recall period also interacts with several other factors. Response effects increase substantially for individuals over 55 as the length of recall increases. Face-to-face interviews typically reduce response errors due to poor recall by reducing omission errors, although face-to-face interviews may also stimulate overreporting. Threatening topics increase response effects for longer recall periods due to omission. Also, open-ended questions have been found to yield somewhat weaker response effects over increasing recall periods than do closed-ended questions.

Saliency of the information requested might be expected to have an important effect on responses. Surprisingly, Sudman and Bradburn found that saliency of questions had no general effect on behavioral

items. However, low saliency increased response effects for attitudinal items. The authors speculated that the difficulty of measuring saliency might have been responsible for the lack of a strong effect.

In short, long recall periods increase response effects. Length of recall interacts with several other task variables in producing this effect. No general effect due to saliency of information was observed, although low saliency increases response effects for attitudinal items.

Method of Administration

Analysis of response effects due to method of administration essentially involves a comparison of face-to-face with self-administration. Those effects are generally best interpreted as arising from interaction with other task characteristics. Most notably, as mentioned above, respondents are far more willing on self-administered interviews to give information placing them in an unfavorable light. In fact, most of the response effects attributable to method of administration probably arise from the fact that face-to-face interviews intensify for the respondent those issues of self-presentation raised by the content of the interview.

INTERVIEW VARIABLES

The biasing effect of interviewer characteristics on survey responding is the most widely researched area in the response effects literature. Far more effort has been devoted to this topic than to the assessment of response effects due to task variables. Nevertheless, Sudman and Bradburn (1974)

conclude that response effects due to interviewer characteristics are generally less important than those due to task variables. When interviewer variables exert important effects, it is usually in the context of highly specific relationships among interviewer characteristics, respondent characteristics, and the topic of the interview.

Sudman and Bradburn also found that the social status of the interviewer could interact with task and respondent variables to engender response effects. Katz (1942) for example, found that working-class interviewers obtained more pro-labor responses, particularly from union members, than did middle-class interviewers. On war issues, middle-class interviewers obtained more interventionist responses than did working-class interviewers. The finding may be of importance in interviewing mentally retarded persons, particularly where the respondent might regard the interviewer as having authority.

Sudman and Bradburn (1974) found no general effect due to sex or race of the interviewer, or to any other combination of interviewer characteristics. The only exception to this general lack of interviewer effects pertains to "don't know" answers. Fragmentary evidence suggests that interviewers over 50 get more "don't know" responses, that interviewers of higher social class and with more experience get fewer "don't know" responses, and that the "don't know" rate may be higher for female interviewers.

Interactions of Interviewer Characteristics with Other Variables

There are some interactions between interviewer characteristics and method of administration,

question structure, and self-presentation issues.

The most notable response effects attributable to interviewer variables occur when there are specific combinations of interviewer and respondent characteristics and topic of interview. For example, Schuman and Converse (1971) conducted 500 interviews in black households, using both black and white interviewers. Response effects due to interviewer race were large for items having a definite racial content, especially items involving militancy or hostility toward whites. In response to the question "Do you personally feel you can trust most white people, some white people, or none at all?", 35% of respondents replied "trust most whites" when interviewed by a white person, as opposed to only 7% when asked by a black interviewer. Schuman and Converse observed little or no effect due to interviewer race on questions not involving militancy or hostility to whites. This illustrates the general conclusion that response effects attributed to interviewer race or sex are highly specific to the interview topic.

In summary, larger response effects tend to occur when the interviewer is young or inexperienced. The most notable response effects occur, however, when the topic of the interview is such that some combination of interviewer and respondent characteristics may bias respondent's answers.

RESPONDENT VARIABLES

It seems likely that some types of respondents are more prone than others to giving biased answers. Sudman and Bradburn (1974) found no general response effects associated with the sex, race, or age of the respondents. They did find that the

largest response effects tend to occur for respondents with eight years or less of school. This result does not refer to poorly educated adults but to children who are still in school. It is not clear what characteristics of school children make them especially susceptible to response effects, but this finding may have important implications for interviewing mentally retarded persons. Mentally retarded individuals are both "young" in the sense of mental age and uneducated, in that they do not progress normally through the educational system. Thus, if it is either low mental age or lack of education that exaggerates response effects in school children, we may expect that mentally retarded individuals will be particularly susceptible to response effects.

Sudman and Bradburn also report some response effects due to particular combinations of respondent and task variables. Respondent variables interact to a minor extent with level of threat, method of administration, structure of questions, length of questions, and the possibility of a socially desirable answer. Two respondent-task-interactions, already discussed under task variables, are of particular relevance to our research. First, response effects for closed-ended questions (i.e., over-reporting) are particularly high for elementary school students. Second, various response effects engendered by long questions are particularly strong for respondents with high school education or less.

In short, elementary school students are especially susceptible to response effects. This effect is increased by the use of closed-ended questions. Further, less educated respondents tend to be more strongly affected by long questions.

Implications for Interviewing Retarded Persons

What do these two very different bodies of literature, one on language and communication skills of the mentally retarded and the other on response effects in survey research, tell us about what to expect when we interview retarded individuals? Although we have discussed specific implications along the way, we would like to emphasize the following observations here:

1. There has been almost no research, prior to ours, specifically examining the performance of retarded persons in interview situations. Thus it must be recognized that the literature we have reviewed here is only indirectly relevant to our questions of interest and that any expectations based on that literature may or may not be borne out when retarded individuals are actually interviewed.
2. As a general rule, we can expect retarded persons to behave linguistically much like younger persons of normal intellectual functioning. For the most part, they can be expected to perform like children of comparable mental age. However, at times they may perform below the level one would expect on the basis of mental age, and occasionally (e.g., with respect to vocabulary size) they may exceed such expectations. Such delays in language development represent quantitative differences between retarded and normal-IQ individuals, and thus no need to shape interviews in accord with unique characteristics of retarded persons.
3. Because language development is generally related to mental age, little can be expected of retarded adults who are profoundly retarded or retarded children who have very low mental ages. Certainly interviewing would appear to be unfeasible with persons whose mental ages are below three; on the other hand, answering questions appears to be well within the cognitive and linguistic capacities of those with mental ages above seven or eight. While mental age can be viewed as the most important consideration, other factors may also influence how well an individual can handle the demands of an interview. For example, expectations may be somewhat lower for Down's Syndrome children and for those who are institutionalized.
4. Another factor with major implications for interviewing retarded persons is their high rate of speech and hearing disorders. Even when mental age is relatively high, speech defects and disorders may make it difficult for the interviewer to understand what the interviewee is attempting to say. More crucially, the high prevalence of hearing disorders among the retarded suggests that some screening for hearing defects might be necessary or that alternative (nonverbal) communication systems might be necessary with persons who have significant hearing losses.
5. Studies of semantic development among the mentally retarded point to the need for care in wording questions. Most notably, concrete wording should be favored over abstract

wording, and commonly used words should be preferred over infrequently used words. Studies of the cognitive development of young children can serve as a useful guide in deciding what types of concepts might be too cognitively advanced for many retarded persons.

6. Similarly, the structure of questions asked of retarded persons should be kept as simple as possible. Questions with dependent clauses, passive constructions, negatives, and so, because they are transformationally complex, can be expected to be difficult to understand. To be preferred are questions that are based on simple active declarative "kernel" structures and that are as short as possible. Studies of question comprehension and production among normal children suggest a regular developmental sequence that can serve as a guide to constructing questions. For example, yes-no, what, and where questions can be expected to be easiest for retarded persons to understand and answer, while who, how, and why questions, and especially when questions, appear to require a higher level of cognitive development.

7. Retarded individuals, like normal children, can generally be expected to understand more advanced linguistic forms than they can produce on their own. As we have noted, the implication of this for designing interviews is that structured question formats may prove easier to answer than open-ended formats because they reduce the demands on the interviewee to compose and express lengthy

answers. Questions should still be kept as simple as possible to aid comprehension, but anything that would make answering easier seems particularly advisable (e.g., asking respondents to point to a response, give a single word answer like "yes" or "no," or select one of two or three options).

8. Whereas the literature on language and communication among the mentally retarded suggests that they may have difficulty understanding and answering questions, the literature on response effects in survey research with the general population suggests that *what they say* may not be valid. This research has identified numerous ways in which answers are affected by the nature of the interview task or the specific questions asked, characteristics of the interviewer, and characteristics of the respondent. There is no reason to believe that mentally retarded interviewees are immune from such effects. Indeed, there is reason to believe that their answers might be *especially influenced* by such factors as how a question is worded or whether it is closed-ended or open-ended, for less educated members of the general population as well as young children have been shown to be especially susceptible to response effects.

9. The response effects literature does not always suggest that one way of asking a question is preferable to another; often studies only demonstrate discrepancies between the answers that are given. However, this research does suggest the wisdom of avoiding questions that

contain leading or emotionally laden wording; that are unusually long and complex; and that encourage respondents to present themselves in a socially desirable manner. Acquiescence, or the tendency to agree regardless of question content, would appear to be an especially dangerous threat to the validity of answers given by retarded persons.

10. Finally, the literature on response effects in survey research reminds us that any interview is a social situation affected by characteristics and behaviors of both the interviewee and interviewer. The answers given by mentally retarded persons may be

affected not only by their characteristics but also by such factors as the interviewer's age or the amount of experience or sex has. Moreover, as suggested numerous times in the literature, there may be considerable differences in the way that different interviewees respond to characteristics that are especially likely to yield invalid answers. While our research did not examine the effects of interviewer characteristics, we did look carefully at response biases in the answers given by retarded interviewees, attempting to determine how both the way the question is asked and the IQ of the respondent affect the validity of what is said.

METHODS OF THE STUDY

The project was designed to address the following questions about interviewing retarded persons:

1. To what extent can retarded persons respond to questions in an appropriate fashion, and what factors affect their responsiveness?
2. How reliable are such responses, in the sense of being consistent over short periods of time?
3. How valid are such responses, in the sense of being free of systematic biases and agreeing with information provided by parents or caretakers or documented in records?
4. What types of questions appear to optimize responsiveness, reliability, and validity?

These guiding questions were explored through a series of five studies, each focused on a subsample of the

retarded population and each having its own priorities and basic design. After describing the process of questionnaire development, we will provide an overview of the individual studies and of the methods employed in all of them.

Design of Interview Schedules

In early discussion with representatives of the President's Committee on Mental Retardation, the scope of topics to be asked about was delimited. Three major topics were identified: living circumstances, familiarity with and utilization of services, and opportunities for decision-making. As a result, the interview schedules initially constructed typically included questions about basic characteristics of the respondent, major day activities, extent of involvement in other activities, nature of the residential setting, training currently being received, and independence in daily decision-making.

Interview schedules used previously, including ones used in the Center's own studies of deinstitutionalization, were reviewed to generate potentially useful questions and ideas for alternative question phrasing. Pilot interviews were developed during the first part of 1977, and they were tested on a small sample of adults, both institutionalized and living in the local community, in May, 1977. Extensive revisions and analyses of individual items were then conducted, and the pilot interview schedule was sent to PCMR for review. After further refinement, questionnaires were constructed for use in the first interviewing venture with institutionalized children and adults in August, 1977. In these studies, which addressed the reliability question, some questions were simply repeated in exactly the same form on two interview schedules administered six to eight days apart. In addition, in these studies as well as in later ones, a major effort was made to include alternatively structured or worded questions on the same topics in order to test the relative merits of various questioning strategies. For example, in asking about decision-making, we tested these two alternatives: "Who decides what clothes you put on in the morning?" versus "Is it up to you what clothes you put on in the morning?"

It should be noted that while some effort went into establishing the content of the interview schedules, it was minor in comparison to the effort that went into developing workable questions and alternative question forms. In other words, we were not overly concerned with developing a standard protocol to fully address the questions of living circumstances, services, and

decision-making. Instead, while using these topic areas as a context, we concerned ourselves primarily with exploring feasible interviewing strategies. On occasion, we included questions which we expected to fail simply to document that certain approaches are relatively unworkable. On occasion, too, we deviated from the three major topic areas in our quest for understanding of interview behavior. Finally, we did a considerable amount of alteration of questions from study to study, refining questions that had been used before and introducing new questions appropriate to the samples under study, so that, in the end, the various samples we interviewed responded to very few of the same questions. Appendix A, by presenting many of the questions that were included in the various interview schedules, offers a more concrete illustration of the kinds of questions which were asked.

The Five Study Samples

Because the study was exploratory in nature, its design evolved over time, but from the start there was an interest in testing the feasibility of interviewing with different segments of the mentally retarded population. The five studies collectively involved both children and adults, in both institutional and community settings, as well as identified significant others (parents, attendants, and advocates). In all, 180 mentally retarded persons were interviewed. Since some participants were administered repeated interviews, the total number of client interviews held was close to 320, and in addition, 210 interviews were conducted with significant others. The five studies and their samples are described on the following page.

INSTITUTION CHILDREN

The study of institutionalized children focused on 52 children, age 12-16, who were residents of a single state institution for the retarded: 20 severely retarded by Stanford-Binet cutoffs (IQ 20-35), 16 moderately retarded (IQ 36-51), and 16 mildly retarded (IQ 52-68). A list of eligible participants was drawn by random sampling of the facility's files, and provision of consent and availability for interview then determined which eligible subjects (those in the right age and IQ range) were actually interviewed. Two interview forms were used, with some questions repeated and others phrased differently on the two forms, and with the forms administered approximately a week apart. Through counter-balancing, half of the children received one form first and half received the other, half were interviewed by one interviewer first and half by the second, and each child was then administered the other form by the other interviewer, as shown in Table 3.1. A cottage attendant familiar with each subject was also interviewed using an interview schedule which included most of the same questions asked of residents (with "you" changed to "he" or "she").

INSTITUTION ADULTS

The study of institutionalized adults was identical in design, but involved adult residents of the same institution, including profoundly retarded persons. A total of 58 adult subjects were ultimately interviewed: 16 profoundly retarded, 16 severely retarded, 16 moderately retarded, and 10 mildly retarded. The interview schedules used with this group were longer and more intricate than those used with any other groups, and again attendants were asked most of the same questions.

COMMUNITY CHILDREN

The third study focused on children in the 12-16 age range; 57 in all, 18 severely retarded, 19 moderately retarded, and 20 mildly retarded; none of whom were institutionalized. Use of the 12-16 age range in both this sample and the institution children sample made them comparable in approximate mental age (the five- to eight-year old range). Students at the lower IQ levels were drawn from a school for the trainable mentally retarded, while mildly retarded subjects were drawn from other special education programs in the same city. Informed consent from parents, obtained by

Table 3.1: Study Design

1st Administration	2nd Administration
1/2 of sample: Form A, Interviewer 1	Form B, Interviewer 2
1/2 of sample: Form A, Interviewer 2	Form B, Interviewer 1
1/2 of sample: Form B, Interviewer 1	Form A, Interviewer 2
1/2 of sample: Form B, Interviewer 2	Form A, Interviewer 1

home delivery of forms by children, determined which eligible subjects were in the final sample, thus making it impossible to determine the extent to which the sample was representative. This study did not deal with the interviewer-reinterview reliability issue, but instead tested several new questions appropriate to community residents and several alternative question formats on the same topic. Each subject's parent or guardian was also interviewed.

The original design of the project called for extending the research to a fourth major group of interest: adults in the community. However, by the time that final phase was being planned, we had had an opportunity to examine data from the first three samples and decided to alter the plan. It had become apparent that interviewing severely retarded persons and interviewing more highly verbal, mildly retarded persons were two different undertakings. On the one hand, we still wished to interview adults living in the community but felt that it might be possible with such a group to try out more unstructured questions and complex topics. On the other hand, we felt a need to search further for techniques which could be successfully employed with low verbal persons. As a result, two additional studies, far different in samples and methods were conducted.

COMMUNITY ADULTS

To facilitate identifying a diverse sample of adults living in the community, we made contact with the Association for Retarded Citizens in another city in Texas and requested their cooperation in the interviewing study. Letters and informed consent forms were mailed to ARC-based citizen advocates, who were to indicate whether they would be interested in being interviewed themselves. They

were also asked to explain the project to their proteges and determine if they too would be willing to be interviewed. Unfortunately, of the 70 some letters mailed, very few were returned, and even with follow-up phone calls informed consent was obtained and interviews could be scheduled with only 13 mentally retarded participants and their 13 advocates. The retarded participants, whose exact IQ scores were not obtained, were generally in the mild or moderate range of retardation and represented a highly verbal group. Interviews with them were conducted by a single interviewer and were tape recorded. The study, which yielded long and enlightening transcripts, provided an opportunity to use unstructured probing techniques and to explore difficult areas such as satisfaction with services and reactions to being, or being considered retarded. While we will not provide detailed findings in view of the small sample size, we will occasionally quote from the transcripts to illustrate processes and problems in interviewing.

REINTERVIEWING OF LOW VERBAL CLIENTS

The final study in the series was designed to further explore techniques of eliciting information from low verbal subjects. It involved reinterviewing institutionalized children and adults who had been largely, but not totally, unresponsive in the previous institution studies. The 29 subjects ranged in IQ from 20 to 48 and were in the lower half of the distribution of responsiveness, but had been able to answer at least one question in the first interviews. The phenomenon of acquiescence was explored further in this study and new approaches using pictures were tested out. Characteristics of the samples are summarized in Table 3.2.

Access to Subjects

Since gaining permission to conduct a study and gaining consent from potential participants are part of any interviewing study, a word on the procedures used in the present study is in order. We elected to make contact with particular agencies or facilities, each of which had its own procedures which we were then obligated to follow. If we had been seeking a representative sample of mentally retarded persons of all ages and circumstances, the procedures used would have been far more elaborate and difficult to implement

than the procedures we actually used. To compose the institution samples, an approach was made to a facility in Arkansas with which the Center had had previous relationships. A proposal was submitted for review and approval by the body charged with determining the appropriateness of research proposals affecting clients in the Arkansas MR Services system. Although it did not at first appear that consent from all parents would be necessary, a decision was made that such consent would be necessary. As a result, a trip that was to have been the data collection trip turned

Table 3.2: Characteristics of Samples

	IQ		Age	
	M	SD	M	SD
Institution Children (N = 52)	42.08	13.81	14.81	1.25
Institution adults (N = 58)*	39.76	13.12	23.49	4.42
Community children (N = 57)	47.53	12.12	14.03	1.92
Community adults (N = 13)				
Reinterviewed institution sample (N = 29)	31.62	7.24	19.58	4.76

*The entire adult sample consisted of 58 persons, 16 of them with IQs from 0-19. However, as they were generally uninterviewable, and as most analysis focuses on the severe through mild ranges of retardation, they are not included in these figures, which are instead based on 42 participants in the severe to mild range of retardation.

into a lengthy process of gaining consent. Parents of children and adults potentially eligible for the study were called by project staff members in the presence of a witness and asked to provide their consent. Following this, letters were sent to those parents or guardians who had given consent verbally so that their consent could be provided in writing. In addition, each client's verbal consent was obtained before the interview was actually begun.

In the study of community children, the planned research again had to be approved through normal mechanisms in the school system. In this case, there was a prohibition against the research project staff making direct contact with parents or children, so a designated staff member in the central administration was responsible for distributing explanation letters and consent forms to teachers of mentally retarded students in the specified age range and then collecting the forms which found their way home with the students and were returned. Using this procedure, many forms were never returned, and when it became apparent that few mildly retarded subjects were being identified, new rounds of forms were distributed to teachers of the mildly retarded.

Similarly, to constitute a community adult sample, we first had to obtain the endorsement of a local Association of Retarded Citizens. In this case, a project staff member was given permission to mail letters and forms directly to all citizen advocates and to directly receive the returns, and written client consent was also required. Again, many forms were not returned. In no study was the rate of refusal high, but in those that involved return mailing of consent forms, low rate of return was the general rule. Unfortunately, this creates uncertainty about how representative those who actively

provide consent are. The telephoning procedure used to constitute the institution samples was highly effective by comparison, but in view of the recent tightening of procedures for access to subjects, it may not continue to be acceptable, for it does involve a facility's making available to outside researchers clients' names and addresses. The moral is that an interviewing project of this nature must leave adequate time for obtaining permission to conduct the study and for gaining consent. There are special problems in making the nature of the study clear to mentally retarded people. Appendix B discusses some of the current issues in obtaining research approval and informed consent. Appendix C illustrates how the present study was explained to mentally retarded persons immediately before they participated and includes a sample consent letter to parents.

Measuring Responsiveness

An important focus of the project throughout was the ability of mentally retarded persons to answer questions in a manner appropriate to the form and content of the question, whether or not the answer was "true" or valid. *Responsiveness* to the communication demands of questions was measured in each study through a slight modification of a coding system originally developed by Sigelman and Werder (1975), which included nine categories:

1. no response
2. unintelligible response
3. irrelevant response
4. don't know, don't remember, not sure (when respondent can be presumed to know)
5. inadequate response (vague or uncodable)

6. request for clarification of the question (huh? response)
7. refusal to answer
8. minimally appropriate response (response which meets the formal demands of the question by providing one relevant bit of information, e.g., "yes" or naming one item in response to an open-ended question)
9. expanded response (response which provides qualifying or additional information beyond the formal demands of the question)

Interrater reliability for this coding system has consistently been established at .85 or higher. In all studies, the interviewee's first response to the question was coded in this manner by the interviewer either during the interview or afterwards using notes on the response given. In the adult institutional sample, responses to a second asking of the question when the first response was inappropriate were also coded in this manner. In the children's institutional sample questions were asked only once, while in all other samples questions were asked a second time, if necessary, in exactly the same form. While responsiveness to the first asking formed the measure of responsiveness in all studies, codings of the contents of responses were based on whichever response, first or second, was at least minimally appropriate. The number of responses coded 8 or 9 (minimal or expanded) was divided by the number of questions asked to form individual responsiveness scores. To form difficulty scores for given questions, the number of interviewees whose responses were coded 8 or 9 was divided by the number of persons asked the question.

In all studies a simple screening procedure was used when a subject's ability to engage in an interview was suspect. The procedure consisted of three

simple questions ("What is your name?", "Is your name (name)?", and "Can you understand me?"), each of which was asked a second time if no initial response was obtained. If any response (word, sound, gesture) to one of these questions was given, a subject was administered the full interview. Those who would not respond at all were defined as totally unresponsive to all questions in the full interview, even though it was not actually administered.

Analyzing Reliability and Validity

In all studies contingency tables were analyzed to examine issues of reliability and validity. In comparisons of subjects' responses to a first and second asking of the same question or to alternatively phrased questions on the same topic, contingency tables indicated the percentage of the responding sample providing the same answers on both occasions. Similarly, subjects' responses can be cross-tabulated with significant others' responses to the same questions. At the same time, cells representing inconsistency between paired responses can be examined to identify particular patterns of response error. The sample contingency table presented in Table 3:3 will give an illustration of the method used. It presents proportions (rather than actual numbers) of various pairings of client and significant other responses about church attendance. If we look at the marginals of the table, we would note that there are discrepancies between client and significant other responses, for 80% of the clients claim that they go to church while 70% of the significant others claim that clients go to church. If we add the proportions of yes-yes and no-no combinations of response, we have an overall agreement figure indicating that in 70% of the cases in which answers were paired, client and significant other agree. Examining the cells representing dis-

Table 3.3: Sample Contingency Table

Significant other: Does client go to church?	Do you go to church?		
	Yes	No	
Yes	.60	.10	.70
No	.20	.10	.30
	.80	.20	

agreement, we note that it is more common for the client to say "yes" and the significant other "no" than for the client to say "no" and the significant other "yes." In this instance, then, we have detected a systematic bias in responses that accounts for disagreement - one which probably reflects the acquiescence phenomenon with which we will concern ourselves later.

On the face of it, 70% agreement sounds fairly good. However, note that both clients and significant others were highly likely to answer "yes" to this question, and that as a result, the odds by chance alone of obtaining many yes-yes combinations of answers is rather high. In other words, in tables like this one, we cannot simply assume that by chance alone client and significant others would have agreed 50% of the time. Unfortunately, traditional statistical techniques designed for use with contingency tables, measures of strength of association such as ϕ and gamma and lambda and measures of statistical significance such as chi-square, do not provide an appropriate means of determining how much a given agreement figure exceeds

what can be expected on the basis of chance. However, Cohen (1960) has described a statistic called *kappa* which appears to suit this problem perfectly.¹

Because the project has involved hundreds of contingency tables, we have not calculated kappas for all tables, but we will make reference at various places in the report to kappa statistics, for they provide us with a means of putting agreement figures in some perspective. The same logic that we have described here also applies if one is concerned with whether clients answer the same question the same way on two occasions and whether their answers to two different forms of question are consistent.

Footnote

¹To illustrate kappa, note that the obtained proportion of agreement for Table 3-3 is .70. The proportion expected by chance is easily calculated from the marginals associated with the two cells representing agreement. The chance probability of a yes-yes combination is .56 (.70 X .80), while the chance probability of no-no combination is .06 (.30 X .20). Therefore the proportion of agreements expected by chance is .56 + .06 or .62, quite a bit higher than 50-50. Kappa is calculated as follows:

$$k = \frac{p \text{ obtained} - p \text{ by chance}}{1 - p \text{ by chance}}$$

In our example,

$$k = \frac{.70 - .62}{.38} = .21$$

Very simply, this means that 21% of the cases represent agreement beyond agreement that one would expect on the basis of chance. Kappa ranges from -1.00 to +1.00, a negative kappa indicating that disagreement is more likely than one would expect, a kappa around zero indicating that agreement is not very different from what would be expected based on chance, and only a large positive kappa indicating agreement beyond chance probability. The statistical significance of a kappa can be determined by first calculating a standard error term,

$$\sigma_{ko} = \frac{P_c}{N(1 - P_c)}$$

or in the example, assuming an N of 50,

$$\sigma_{ko} = \frac{.62}{50(1 - .62)} = .0326 = .18$$

$$\text{then } z = \frac{k}{\sigma_{ko}} = \frac{.21}{.18} = 1.17$$

By referring to a normal curve distribution table, we find that the probability of this z value is .12, approaching the conventional .05 level of probability but not large enough to be considered statistically significant or a systematic departure from chance probabilities.

INTERVIEWING RETARDED PERSONS: LESSONS FROM EXPERIENCE

Most of the subsequent chapters of this monograph concern the interview responses of mentally retarded persons and their parents or caretakers--the data of the project. However, we have gained, through years of experience interviewing mentally retarded persons, another type of "data," often subjective in nature, about the process of conducting interview research with the mentally retarded. Since mentally retarded people are, after all, people, much of what we would have to say is simply a matter of sound survey research design and practice. The literature on survey research is helpful as a source of general principles of interviewing applicable in studies of the mentally retarded. In Chapter 2, for example, we discussed many of the variables such as question wording, interview behavior, and subject characteristics which can produce error in any survey research project. Field interviewer manuals developed for surveys of the general population are also useful. For example, we have found the

Interviewer's Manual developed by the Survey Research Center at the University of Michigan (1969) highly informative about everything from sampling to closing the interview gracefully.

On the other hand, we feel that there are some unique problems in conducting survey research with the mentally retarded. As later chapters will demonstrate, obtaining meaningful answers from low verbal persons is problematic and requires special techniques. In this chapter, we focus on the logistics of interviewing the retarded, drawing on the lessons we have learned through trial and error in hopes that others may be spared our errors.

Designing Interview Schedules for the Retarded

In any survey research project, the researchers presumably have an idea of what they are interested in learning and are guided by this idea in developing questions. The more specific the idea is the better; thus

it is important as a first step to outline in some detail the topic areas of interest, the information goals. It is also important where possible to anticipate which topic areas may be difficult for mentally retarded persons. Much of our data will pinpoint topics about which we had great difficulty obtaining information. There were other topics we tended to avoid from the start because we anticipated problems. Generally, it has been our experience in previous interviewing research that many mentally retarded persons have difficulty with time, number, and money concepts (e.g., giving a chronological account of their whereabouts or their time allocations during the day, or indicating the number of times they have done something in the past month, or describing their incomes and sources of income). Words such as "before," "since," and "after" are difficult (e.g., "Where did you live before that?"). We have also found it difficult to ask about experience with service agencies, as clients often view the agency as "Mrs. Brown" or "Mr. Tom's" and do not have a complete understanding of an agency and its functioning. We have also found it difficult to ask hypothetical questions which ask the respondents to imagine a different world from the one they currently experience. Some of these areas are problematic enough that, where possible, the information is probably best sought from nonretarded persons who know the subject well. In other cases, the process of developing questions will simply be more challenging and time-consuming.

Surveys of the mentally retarded population will almost inevitably involve some branching in the interview schedule; that is, certain responses will dictate that succeeding questions on the topic be skipped. One must know enough about the kinds

of persons one is to interview and their situations to be able to design meaningful subsections and anticipate exceptional cases to whom questions will not apply. However, there is often a risk that the distinctions the researchers think they are making will not be accorded the status of distinctions by retarded persons. For example, when asked if they work, many clients who are in sheltered workshops, work activity centers, or simply helping out their relatives on an informal basis will say "yes." The researcher may have sought a distinction between being competitively employed and being unemployed, but retarded interviewees may make quite a different distinction between work and nonwork. Further problems arise when clients do not distinguish between sheltered workshops, activity centers, and so on. In fact, they may know these types of settings only by their local proper name. Follow-up questions are often needed to make possible the kinds of distinctions that the researchers had in mind.

Similarly, we have encountered many difficulties associated with question phrasing, such that respondents may not be making the same distinctions we choose to make between, for example, being able to do something, doing it at least sometimes, doing it regularly, or being allowed to do it. The more prior knowledge one has of retarded persons and their views of the world, the more likely it is that questions will be sound. However, pilot testing is the best single method of checking to see that the questions are working and determining whether additional follow-up questions might be needed to clarify answers.

There is a natural tendency in the question development phase for questions to proliferate. One naturally seeks more specificity

about any area to be covered, but also tends to become sidetracked in fleshing out an area of questioning. Having clear information goals as a guideline is one check against this tendency, but later editing may still be called for after pilot testing. It is our feeling that the interviews themselves should not consume more than about half an hour or it becomes difficult to maintain active attention.

Our interviewing studies have each required two major forms besides a client interview schedule: a record form and a significant other interview. The record form is used to collect whatever information is available from agency or facility files about demographic characteristics, IQ score, length of institutionalization, daily schedule, and so on. In our studies, parallel interviews with "significant others" (parents, caretakers, or others who know a client well) have typically asked almost all of the same questions asked of retarded persons directly, excluding some subjective questions. This was done in order to explicitly examine the correspondence of answers. In most studies, however, there are some types of information, as we suggested above, that are best gotten from significant others or caretakers for which the interviewer would at least want to seek verification of answers given by retarded persons. In many studies, significant others might also be used as the major source of information about such things as clients' histories, current skill levels, and degrees of independence. Indeed, we find it hard to imagine an interviewing study which would not require at least a minimal amount of data from files or from knowledgeable informants. We found it best to interview such informants personally rather than

have them fill out the questionnaire on their own, unless the questions were very simple and straightforward. The same principles of survey research applicable to retarded persons apply to their significant others, for their answers too are subject to bias, and, in some cases, they lack knowledge about certain areas. This is true even of parents.

Pilot Testing

Assume that careful planning has led to the development of an interview schedule. As noted above, much of the information in the present report is intended to guide the task of designing questions which can be answered reliably and validly by mentally retarded persons, so we will not discuss question design further here. What we want to stress is that what appears to be a good interview schedule may not turn out to be one in practice. Take the simple question, "Do you make your bed?" As our data will indicate, this yes-no question may be subject to acquiescence which inflates the number of interviewees reporting that they make their beds. Moreover, the question is ambiguous. It does not indicate, for example, where bed making takes place. Often institutionalized persons do some things at home that they do not do in the institution and vice versa. The question also leaves open the question of extent of activity since a "yes" response may mean that the respondent has made a bed once or that he or she makes a bed every day. In constructing questions, one must adopt a highly critical attitude, actively looking for ambiguities and actively imagining potential misinterpretations of the question. It is helpful to imagine interviewees in diverse situations to judge whether a question is indeed applicable and meaningful to the whole

range of persons to be interviewed. Instead of thinking about the general rule, one does best thinking about all imaginable exceptions to the general rule.

Imagination can carry one only so far, however. Pilot testing of interview schedules is, in our opinion, essential. A first step may be role playing the interview with co-workers playing retarded persons. This at least provides an estimate of the approximate length of the interview and identifies questions which may have looked good on paper but do not sound good when they are asked. Those playing the role of interviewees can be asked to actively search for chances to misinterpret the question or raise questions about its meaning. The next step should be pilot interviews with retarded persons similar to the ones who will actually be interviewed. Ideally, these pilot interviews should be tape recorded, and the interviewer should feel free to rephrase questions on the spot if they do not seem to be understood. Such rephrasings and the responses to them are often useful in refining questions. The interviewer can also probe to determine if the response given is what the interviewee meant to say in cases where the validity of response seems suspect. We have found that conducting pilot interviews inevitably reveals problem areas that would otherwise go undetected: questions that few can answer, questions that are greeted by puzzled expressions, questions that are quite consistently misinterpreted, and questions or questioning strategies that prove embarrassing for both interviewer and interviewee. In order to gather information about the validity of responses, it would also be useful to compare clients' answers to documented fact or selected responses given by nonretarded informants in parallel interviews.

Interviewer Training

It is our opinion that virtually any sensitive and conscientious person can become an effective interviewer of retarded persons. The more experience the interviewer has had with retarded persons, the better, however. We suspect, although we did not encounter the problem ourselves, that some people, unfamiliar with retarded persons, may not work out as interviewers if they have negative reactions and those reactions are noticed by interviewees. In previous research, we actually encountered similar problems with some persons experienced in mental retardation--usually agency staff who carried an authoritarian attitude into their interviews. Assuming that willing and able persons are recruited, the main tasks are to train them to use the interview schedule and to standardize their behaviors as interviewers.

Interviewers should, of course, be well-versed in the content of the interview schedule before they conduct interviews. A serious piloting effort can provide the setting for training. Where feasible, it is useful to have two interviewers record the answers given by pilot subjects so that the second interviewer can provide feedback on the first's approach and the data collected by the two can be compared for reliability.

Interviewers not only have to be able to use the interview schedule confidently, but they must be well-versed in all aspects of conducting the interview. They should have a common understanding of any question rephrasing or probing procedures to be used; they should have a consistent approach to providing non-specific feedback to interviewees to maintain rapport and encourage answering; and they should establish in advance how to handle various problems that may arise. Research

discussed in Chapter 3 suggests that interviewer characteristics and behaviors can at times influence responses. Certainly if one interviewer is highly warm, supportive, and attentive while another goes through the motions in a cold and mechanical manner, one can expect very different results. There is also the risk that a warm and responsive interviewer can become more mechanical after hours of back-to-back interviews. Thus, interviewer manner can be as important as interview content.

Locating Interviewees and Scheduling Interviews

In some types of studies in which subject's addresses have been obtained from an agency or facility, simply locating the subject and arranging for an interview can be problematic. When one arrives at the supposed address and finds that the client no longer lives there, one can check with neighbors, the telephone and utility companies, the Post Office, or the Police Department. City directories usually on file with the Chamber of Commerce can also be useful. Other potential sources of information include local churches and small neighborhood grocery stores. The interviewer on a search should always explain his or her affiliation and purpose, but should avoid indicating that the persons being sought are mentally retarded. As Edgerton (1967) has vividly shown, for example, many deinstitutionalized persons have struggled to avoid labeling and stigma; they would not appreciate having their "covers" blown by an insensitive interviewer.

Initial contact with an interviewee can be made either by telephone to set up an appointment or by going to the home or residential facility. Often cooperation is more likely if one goes directly to the home,

but telephone contact is more cost-effective and works well if the client has already provided consent by mail. If telephone contact is made to schedule appointments, it will still be necessary to introduce yourself and the study briefly when you arrive for the interview as time will have passed in the interim. Plenty of time for travel and unanticipated delays (often informal chatting with interviewees or their family) should be allowed between interviews.

Obtaining Consent

In Appendixes B and C, we present additional information on research guidelines relevant to survey research as well as sample consent and explanation forms used in the present study. The points below simply summarize the kinds of information needed to make consent *informed* consent. There is a special burden on interviewers of the retarded to insure that the study is understood, and extra effort and patience may be required. The emphasis in communications with retarded persons as well as their parents and care providers should be on simple, concrete, and even purposely redundant language. In obtaining consent for an interview, explain the following in simple, concrete language:

1. Who you are
2. Who the interviewer will be (if it is not you)
3. For whom you are doing this
4. Why you want to talk to them
5. How long the interview will take
6. What kinds of questions you will be asking
7. When you will be conducting the interview (approximately; schedule specifics based on

their schedules and desires as a matter of courtesy)

8. Where it will take place (Typically their own living situations put them most at ease, but it should be left open in case they prefer another place)
9. Who else is involved--e.g., how you got their names, others you will be talking to about them, etc.
10. How it will affect them (or the fact that it will not affect them if all information is to be kept confidential)
11. Who will see their answers and who will receive any reports that are based on the answers of many interviews
12. How you will be asking for their consent and what they need to do to give it
13. That participation is voluntary--that they do not have to be interviewed if they do not want to, that they can end the interview at any time, and that they can skip questions they do not want to answer
14. Whom to contact and how to do so should they have any questions

Depending on the regulations for access to subjects in effect, these points are either presented in a letter requesting consent to participate or they are presented orally in person or by telephone, before the interview is to be conducted. Our preference is for the oral presentation not only because it gives the client the opportunity to ask questions and clarify any misunderstandings, but also because the return rate on mailed consent forms is usually quite low. Moreover,

when written consent is obtained by mail, there is often a long time lag between the giving of consent and contact with the consentor for purposes of scheduling an interview. At that time, a rehash of the important aspects of the study is often required to remind the subject of what was agreed to earlier. If not previously obtained by letter, written consent should be obtained after oral consent is given. After consent is obtained, we reinforce the following points immediately before the interview begins:

1. The interview is not a test, and there are no right or wrong answers.
2. Everything said will be kept a secret (assuming that is indeed the case).
3. The interview will not change their lives in any way. We hope that it will help other people like them someday, but we are not going to do anything to help interviewees with any problems they might have, nor will we make them do something they don't want to do--like move or go to another school.
4. Some of the questions may sound silly, but we need to ask all of them. (Our surveys indeed involved blatantly silly questions as checks on acquiescence, but virtually any survey contains some questions that make interviewees wonder. We use the term "silly" rather than "dumb" or "stupid" because its connotation is more favorable. We also point out other peculiarities, e.g.) the fact that some questions are asked more than once.)
5. They do not have to answer any questions they don't want

to, and can end the interview any time.

6. If they need to stop, to get a drink or go to the restroom, they should feel free to say so.

Interviews with retarded persons do not technically require parent or caretaker consent if the client is of age and has not been judged incompetent. However, it is a reality that parents and others who are responsible for retarded persons have a stake in what happens to them. Whatever consent procedures are used, it is important to make contact with such significant others so they are assured of what you are doing and know that it will not affect the client's life or theirs or involve any risk of harm to the client. It is almost always necessary to make appointments for interviews directly through such persons as you will be in their homes or their residential facilities. Due to the difficulties that some retarded persons have with time concepts, such persons are also likely to be more aware of client's schedules than clients themselves are. The major point, however, is that the position of a person responsible for the retarded interviewee should be respected despite the fact that their formal consent may not be required. When parents and caretakers understand the study, they are put at ease and are more likely to accept your presence. We have encountered the most reluctance in some residential facilities where staff fear the interviews may put the facility in an unfavorable light. In one interviewing study in which we were involved, a whole group of clients at one facility had to be dropped from the sample because houseparents were afraid that the study would "dig up" something. Anticipating these kinds of reactions beforehand, you can be

ready with answers to concerns that might be raised.

Getting Ready to Interview

We would not think it worthy of mention if it had not happened occasionally, but one can never do enough last-minute checking. One of our interviewers, after numerous calls and visits, finally lined up an appointment for an interview, only to arrive, open her briefcase, and find that she did not bring the correct interview schedule with her. Pencils break, mimeographed forms sometimes have missing pages, and whatever can go wrong often does.

When one is interviewing the retarded in their homes or in residential facilities, it is important to try to find a quiet place away from others in the household. We have repeatedly encountered difficulties when parents or caretakers attempt to take over and answer questions for the clients or react to what they are saying. The best way to avoid this is to explain the need for privacy and your interest in hearing what the client has to say, even if he or she may be giving factually incorrect information. Sometimes it is helpful just to tell the parent that many clients will want or expect their parents to answer some of the questions and that you need to avoid this if possible. Also, explaining that the same questions will be asked of them, if that is being done, alleviates their fears of the client giving out misinformation. If you intend to give a questionnaire rather than an interview to the significant others, it may be given to them to fill out while you are interviewing the clients. This will keep them occupied so they are not so likely to be involved in the client's interviews. When the household is particularly noisy and crowded or otherwise inimical to a successful interview, we have sometimes asked the client if he or she

would mind sitting outside for the interview. Although parents must be allowed to sit in if they insist, they may need to be reminded to let the client answer the questions. This arrangement is not as desirable as total privacy, but it is sometimes unavoidable.

Relating to the Interviewee

Although obtaining consent to be interviewed is obviously important, it is perhaps more important to move beyond mere consent and attempt to make the interview an enjoyable experience. An interview should be a give and take situation. In other words, instead of focusing exclusively on one's own need to obtain information, be sensitive to the interviewee's needs. We have discovered that many mentally retarded persons find the experience of having someone sit down individually with them and listen to them rare and gratifying. They may want you to look at their pictures, listen to them talk about something unrelated to the interview, or meet their friends. Often you can suggest that such things be done as soon as the interview is concluded, and this will satisfy them as well as provide an incentive for continuing the interview. At other times it is best, particularly when clients want to tell you something important, to take a break on the spot and let them talk. This is why it is necessary to schedule more time than is required to do the actual interview. If an interviewee persistently wants to pull you off the course, politely steer him or her back to the questions, with something like, "That's interesting. Now I'd like to ask another question."

On occasion, a mentally retarded person will display inappropriate behavior in an interview. Our interviewers have encountered the gamut, from masturbation, to pulling

the hair of the interviewer, to getting up and wandering around the room, to simply carrying on a monologue totally unrelated to the interview. One of our interviewers simply ignored the fact that her interviewee was masturbating during the interview, but not all interviewers will find this quite sound approach of ignoring inappropriate behavior acceptable. If, as occasionally happens, a client poses a physical danger to the interviewer or to himself, terminating the interview is the obvious course of action. In response to most disruptive behavior, the best policy is to be firm but polite in attempting to steer the client back to the interview as well as to reinforce, through smiling and words of encouragement, paying attention and answering questions.

Perhaps the most important word in interviewing retarded persons is respect. Throughout history, mentally retarded persons have been treated as less than fully human individuals. Repeatedly they find themselves in situations where two nonretarded people talk about them as though they were not there instead of asking their opinions directly. In scheduling an interview, in conducting it, in closing it, the interviewer should remember that the interviewee's needs come first, that he or she is helping you. It is not enough to think, "I'm going to help retarded persons through these interviews and I'll make them feel that their contribution is important." There is a major difference between this attitude and feeling that the interviewee actually is important and can help you. Retarded persons can detect this difference and are likely to respond much more openly to the extent that you give them reason to trust you and come across as a person who is genuinely interested in them and their opinions, whatever they are.

For some mentally retarded persons, especially those who have left institutions for the community or whose present situations are in jeopardy, there may be a threat implied by the interview. In particular, formerly institutionalized persons and their guardians may worry that their answers will somehow result in their being returned to the institution or being forced into other undesirable living, work, or school settings. They may perceive the interviewer as an adult with power over their lives. It is therefore essential to emphasize strongly both the confidentiality of the information they will be providing and the fact that being interviewed will not change their lives in any way. Of course, if the project is such that it may affect their lives, they should be informed honestly of this during the initial explanation preceding consent. Hopefully your honesty with them will help put them at ease and encourage honest answers.

Sensitivity to the interviewee's needs during the scheduling and consent process is a start, but trust must also be nurtured throughout the interview as well. One way to do this is to provide nonspecific positive feedback along the way. It is dangerous to praise or endorse specific responses during the interview because you may inadvertently shape future responses. To provide nonspecific feedback, you might say, "That's really interesting," or "I'm glad you told me that." Smiles and nods after responses are even less specific and should be used frequently. You should generally have an interested look on your face and avoid any expression of disapproval of what is said. It is best to put any questions that can be considered sensitive near the end of the interview in hopes that by then a relationship of trust will have been fostered. How-

ever, avoid ending the interview on a depressing note, by returning to less sensitive questions at the very end.

Questioning Approaches

Since many retarded persons have both limited verbal skills and various speech defects, simply understanding what they are saying can be a challenge, and there are undoubtedly differences between interviewers in the extent to which they can decode difficult-to-understand answers. Parents and others who know the client well may be able to understand much more and can translate for the interviewer, although as we have noted, the presence of such persons also has definite disadvantages because clients may not feel comfortable saying certain things in their presence. Also, parents in particular may take over the task of answering. We still prefer that interviews be private, and suggest that the interviewer elicit repetitions of answers when they are not understood. With more verbal clients, it may be enough to ask, "What? I didn't hear," but with less verbal ones it is often necessary to actually reask the question ("Let me ask that again; I couldn't get your answer.") Although it is frustrating for both the interviewee and interviewer to go through this process repeatedly, patience may pay off. As long as it is done politely, most clients will try harder to get their messages through.

Let us illustrate the virtue of patience. One adult living in the community was asked, "What things do you not like about living here?" She said an unintelligible word twice; then the interviewer asked, "The what?", and she said the same unintelligible word twice more. After another round, the interviewee finally pointed up to the sky and gestured to indicate that she was talking about noise from airplanes.

(She lived near a busy airport). This woman has a serious speech defect; thus there was a continual need to reask questions in order to understand what she was trying to say and a continual resorting on her part to gestures. It was indeed a frustrating interview for both parties and it did not yield much in the way of information, but at least the interviewer's patience and persistence gave the respondent the fullest opportunity to speak for herself. It is certainly preferable to go through this reasking process to obtain clarification rather than to jump to an interpretation of something said which may be completely wrong.

It may be tempting to check out your hypotheses by saying, for example, "Did you say you like to go swimming?" Unfortunately, since we have discovered acquiescence to be a problem in interviewing retarded persons, this may not provide verification of the response. Some clients may say "yes" simply to be accomodating. This technique is potentially usable with some clients if you have reason to believe that acquiescence will not be a problem, but we strongly favor reasking the question as the first strategy when responses are not intelligible or relevant. In our own research we generally reasked questions only once, but additional reaskings might be an even better policy. Then, if that fails, a good strategy to use when it is necessary to get clarification is to ask an either-or question. This often works even when neither choice is correct. They are more likely to correct your wrong assumption if the probe is structured as an either-or question than if it has a yes-no structure.

In view of findings to be presented later on the effects of question phrasing and wording on the types of answers given, we would also em-

phasize the need to stick with the interview schedule even when clients do not appear to be able to respond appropriately to a question. If interviewers are given the latitude to rephrase questions if the original question does not yield an answer, they have a natural tendency, as do adults speaking to children, to simplify the question, often by converting a difficult question format into an easier yes-no question. If the interviewer feels it is necessary to clarify a response, the question can be rephrased, but the alternatives should be presented in the same form as in the original question. Sample 4.1 illustrates this technique applied to a multiple choice question. As our data will show, this can substantially affect the responses that are given. Similarly, rewording the question, even without changing its format, may in effect change the question. Even a change in emphasis on certain words can change responses (e.g. "Are you usually by yourself?" vs. "Are you usually by yourself?"). The first will elicit responses based on present feelings more than the second because the word "usually" may not be processed unless it is emphasized. In short, if alternative ways of asking the question are carefully planned in advance to minimize any potential for altering the intent of the original question, rephrasing might prove valuable. Generally, however, we favor giving exactly the same stimulus to each person interviewed, and simply treating as missing data cases in which an answer to the question could not be provided.

Also, when presenting multiple alternatives it is important to present all of them. Do not take a subject's response unless he or she has heard or seen all of the alternatives. Often, subjects will try to interject an acquiescent response after each

Sample 4.1

Rephrasing a Multiple Choice Question without Leading the Respondent

The following is an example of how a multiple choice question might be rephrased to obtain an appropriate response without changing the format or intent of the question. The subject here had already indicated that he watches TV.

I: Do you watch TV a lot, some, or not much?

S: Huh?

I: Do you watch TV a lot, some, or not much?

S: Most every night

I: Would you say that's a lot, some, or not much?

S: When "The Hulk" is on. Sometimes I listen to Elvis records.

I: I need you to pick one of the answers here. (This was said in a friendly manner, so the subject would not feel intimidated.) Do you watch TV a lot, some, or not much? (Vocal emphasis was put on all three alternatives equally.)

S: I'll say sometimes. Yeah, some.

Contrast the rephrasing techniques used above with those found in the following example. Note how the interviewer leads the subject into a response.

I: Do you watch TV a lot, some, or not much?

S: Every night I can.

I: Is that a lot?

S: Yeah.

alternative is presented. To avoid this, do not pause between alternative, even if a response is interjected. You will probably have to repeat the question, but by doing this on the first asking, you have indicated that a yes-no response is not appropriate. Most subjects will wait for the full presentation on the second asking. Non-verbal cues are also beneficial. Repeat the phrase while you are looking at the interview schedule. When all alternatives have been given, then look at the client to indicate you want a response.

Similar logic applies to the issue of probing in response to interviewee responses to seek additional information. In some interview studies where in-depth material is sought, this kind of probing can be invaluable because it allows the interviewer to build on what the client says to construct a fuller picture. However, we feel that such probing techniques should also be worked out in advance. We would again warn against the use of yes-no questions as probes, and would favor either-or questions and open-ended questions in order to minimize the chances of leading the respondent in a direction he or she may not have intended to go. Similarly it is dangerous to use reflective techniques of encouraging interviewees to talk more (e.g., "You didn't like that houseparent, did you?") unless you are completely confident of your understanding. Again, less specific types of reflective feedback, simply saying "uh huh," "oh, I see," and the like, are preferable. The transcript in Sample 4.2 illustrates an interviewer's use of probes to clarify a response without leading the respondent.

There are often occasions in interviewing retarded persons when it becomes very difficult to separate fact from fancy. Sometimes, what is being

said simply does not sound plausible. For example we interviewed a man who claimed that he was married to the person assigned as his citizen advocate and no amount of probing would dissuade him from that story. Another adult living in the community was asked the question, "Is anybody teaching you about using money now?" She replied, "I did have, but she was working for the state but she charged money and I couldn't afford to pay." Later, when she was asked if she received help from the vocational rehabilitation office, she returned to this theme, but even through additional probing it proved impossible for the interviewer to clearly establish who this person might have been, for the closest the interviewee came was to describe her as a social worker and relate how the woman did not like the client's barking dog but still "begged" to come every week. She finally told the woman not to come: "I said, 'Take my money? No thank you.' I said I ain't got that much money to pay everybody comes to visit me." Because clients so often do not know what agencies people represent or forget their names, it is very difficult to follow such a story and determine if there is any factual basis to it, even though the story itself is otherwise rich in detail. Verification with someone knowledgeable about the interviewee would be useful, but this cannot always be done. When you have clear evidence that what the client is saying is wrong, you can eliminate the data. On the other hand, one must avoid assuming that a client is wrong, just because he or she describes such things as exploitation or mistreatment by service providers, and one should be sensitive to the feelings underlying even fabricated stories.

Sample 4.2

Example of Probing Techniques

In one of our studies involving unstructured interviews with adults living in the community, the interviewer was free to pursue answers to structured questions through probes, but attempted to do so without leading the interviewee, as illustrated in this example, where such probing was essential to establishing whom the interviewee was talking about. In response to a question asking who her best friend was, the interviewee said "Lois" and then after three structured questions about the friend, the interviewer asked, "How did you meet Lois?"

S: The lady I used to have was named Donna, and she. I don't know how to say it. She got me to know her and that's how.

I: Now who is Donna? How did you know her?

S: My counselor.

I: Was Donna your counselor, or did your counselor introduce you to Donna?

S: She was my counselor and she introduced me to Lois.

I: O.K. Where does Lois work.

S: MMMm

I: Is she just a friend, or does she help you?

S: She helps. She trying to help me find a job, too.

I: O.K. Is Lois a counselor, or is she just a friend?

S: A counselor

I: Do you know who she's a counselor for?

S: It's on ____ Road, the one I know, the place on ____ Road.

(With the help of the answer to a subsequent question, it is firmly established that Lois was a rehabilitation counselor.)

Closing the Interview

After the official interview is completed, we ask if the respondent (1) has anything to add or wants to talk about anything else, or (2) has any questions he or she would like to ask. It seems only fair, since they have helped us, to give them a chance to use our time as they wish. Sometimes, if interviews have delved into sensitive or bothersome areas, the client may want to continue talking about these areas and this may be an important opportunity to work out any unresolved tensions. If you know that a client became upset about some topic raised in the interview, it may be helpful to allude back to it afterwards and give the interviewee a chance to let off more steam. This is also often the time to let the client show you something or introduce you to someone.

In our experience, informal chatting at the end of the interview often reveals more insights than the interview itself. Interviewees are more relaxed because what they say is "off the record." In some cases, they may express negative emotions which they were afraid to express while they were "on the record." It is, of course, unethical to surreptitiously tape such comments, but you may ask to leave your tape recorder on if this information is to be collected. Writing down what they say puts the situation "on the record" again and discourages openness. Another possibility is to take notes after you have left, particularly if what is said adds a different perspective to what was gained in the official interview. We are not sure what to do with such information when it clearly contradicts responses given in the interview. In a small-scale study, one might actually revise answers in the interview on the basis of information surfacing in post-interview discussion, but there are dangers in doing so. It may be best to leave

answers as they were but draw on post-interview comments to enrich one's understanding of the perspectives of retarded persons or prepare in-depth case studies.

Our main point is that you as the interviewer owe the interviewee some of your time if he or she wants it. (Sometimes they do not, and the task is one of recognizing that even if they are too polite to say so.) It is rude and possibly disturbing to the interviewee to make a hasty exit just as soon as you have gotten what you wanted out of the relationship. We prefer viewing the interviewer-interviewee relationship as a human relationship and would like to think that the interviewer leaves in such a manner that he or she would readily be welcomed back. Profuse thanks to everyone involved for their time and help are obviously called for.

Recording Interview Data

In our research, interviewers have attempted to write verbatim responses as they are given. Although tape recording of interviews can be useful, it has two major disadvantages. First, it can be a distraction or source of anxiety for some subjects and may make them less likely to be fully honest. This may be alleviated somewhat by explaining that the tapes will not be heard by anyone else, so that you're taping the interview so you can listen to it yourself to make sure you got everything that was said. Second, transcription of tapes is extremely time-consuming and difficult, especially when the quality is poor due to background noise in residential settings or when the interviewees have speech difficulties or impediments. In studies involving in-depth and largely unstructured interviews, there is probably no alternative to taping, but we would still recommend that the interviewer take as many de-

tailed notes on the spot as possible and use the notes as the primary source of data. Understanding at the time of the interview is typically greater than understanding days later when listening to a low quality tape, partly because mouth movements and body language are aids to understanding. In short, we would recommend use of tapes as back-ups but not as primary data sources. We ourselves used tape in this fashion in conducting relatively unstructured interviews with highly verbal adults living in the community.

Otherwise, we have found it feasible to write down verbatim responses as most retarded persons are not highly verbal. If they do launch into irrelevant monologues, these responses do not need to be recorded. Writing does interfere to an extent with making eye contact and maintaining rapport, so it is important to bear that in mind and attempt to maintain eye contact as much as possible. Occasionally we have encountered another unusual problem. A few clients will stop talking when they see that the interviewer has stopped writing or has run out of room on the interview schedule to write a response. This is a minor problem that can be prevented by leaving large amounts of space after questions likely to elicit long answers. Nonetheless, it illustrates the fact that a host of unanticipated variables affect behavior in an interview situation and the more that can be anticipated and controlled the better.

Finally, we recommend that the interviewer go over the written notes immediately after the interview to make sure that nothing important was incorrectly or uninterpretablely written down. This is essential since memory fades rapidly.

Coding and Preparation for Analysis

To the extent possible, categories of response should be planned

in advance of data collection so that interviewers can simply mark response codes during the interview rather than constructing them on the basis of written accounts of responses. This is quite feasible if the interview is highly structured and uses yes-no, true-false, either-or, and multiple choice formats. It is also feasible with many open-ended questions in areas where some previous research has been done and one can anticipate what the likely response categories will be or has a clear idea of how a universe of responses should most logically be carved up. Typically, however, there are some open-ended questions for which categories of response cannot easily be preplanned. When in doubt, it may even be preferable to wait until the responses are collected rather than to use preplanned codes and later find them to be inappropriate. On the other hand, our experiences in developing coding categories for open-ended questions given by mentally retarded persons have been frustrating. As noted in Chapter 5, such responses are diverse and often strange, and the category schemes evolved are sometimes less than aesthetically pleasing.

In forming categories of response, one aims for a delicate balance between the general and the specific. We have typically used all responses to formulate categories, but in a large study a sample of responses would suffice. In anything, it is better to err on the side of too much specificity. Categories can always be collapsed later if only general categories are of interest, but general categories cannot as easily be broken down again into more specific ones. Inevitably studies of this nature involve some recoding after categories have initially been developed.

Intercoder reliability may be a problem with some questions. We have usually established such reliability informally by having interviewers, after they have attempted to establish a set of coding categories based on similarities of response, explicitly draw up common conventions they will use in deciding which code to apply. We have also done small-scale intercoder reliability checks on two or three interviews to insure that two different coders do indeed code the same responses the same way. Much depends on the specificity of the coding categories. It is straightforward to code a response to a simple factual question as "yes" or "no," but open-ended questions often leave room for subjective interpretation and some type of reliability check should be conducted.

To facilitate preparation of interview data for computer analysis, two approaches seem particularly attractive. In our own research, we used interview schedules which included in the right-hand column variables and computer column numbers (see

example in Table 4.1). The interviewer can either fill in the code on the spot, or, as was the case in search, due to the fact that of both responsiveness and content of response were done, fill it in later. Several blank columns can be left for future use once content categories for open-ended questions have been developed. The advantage is that a keypuncher can punch directly from the interview schedules, and the step of transferring responses to computer coding sheets can be bypassed.

Alternatively, new technologies involving optical scanning can be used to bypass keypunching and the inevitable errors it entails. Response codes can be entered on specially printed optical scanner sheets, entered into the computer, and then analyzed in any manner in which data entered by card can be analyzed. Such sheets can be printed to the researcher's specifications. This approach is especially useful with questionnaire data or tests because respondents can supply their own responses directly on the scanner

Table 4.1: Coding System Used On Interview Schedules

<p>1. Are you usually happy or sad?</p> <p>Sad = 1 Happy = 2</p>	<p>HS C : $\frac{1}{11}$:</p> <p>HS1 R : $\frac{8}{12}$:</p>
<p>2. Are you usually with other people?</p> <p>No = 1 Yes = 2</p> <p>(alone) (with others)</p>	<p>WOP C : $\frac{13}{14}$:</p> <p>WOP1 R : $\frac{14}{14}$:</p>

The blanks in the right-hand column permit keypunching directly from the actual interview schedules. The respondent answered the first question "sad." The "C" in the right-hand column indicates content coding, and the "sad" response is accordingly coded 1. The "R" indicates a responsiveness code; the 8 represents a minimally appropriate response. The numbers beneath these coding blanks indicate the correct column on the cards used in keypunching.

sheets. It would not be feasible to print an entire interview schedule on such sheets, but this approach might reduce the number of errors made in translating data into analyzable form, for the interviewer could complete the sheets directly from the interview protocol.

Based on one experience, we would recommend against the use of what is called free format coding in which the keypuncher works directly from interview schedules and uses a special notation system which does not involve assigning data to card columns during punching but rather doing so later. While this technique may make it easy to transfer data from interview schedules to computer cards, it allows room for many keypunching errors and involves a number of technical problems in then converting the free format data to an analyzable form. Because it demands more technical sophistication than other approaches and because we once experienced numerous problems with it, we do not recommend it.

Conclusion

In this chapter, we have attempted to share some lessons we have

learned about the process of interviewing retarded persons. Unlike subsequent chapters, this chapter is based on subjective impressions rather than hard data. As we have suggested, most of the guidelines for interviewing members of the general population apply equally well to interviewing retarded persons. We have attempted to supplement those guidelines to prepare interviewers for some of the special problems they may encounter in interacting with retarded persons. While we have tended to emphasize problems, we should point out that the majority of persons we have interviewed have been eager to participate, extremely cooperative, and very gratified by the experience of having someone seriously listen to what they have to say. Similarly, most parents and houseparents have been cooperative, often because they hope that research findings will result in improvements in the service delivery system.

TO WHAT EXTENT CAN MENTALLY RETARDED PERSONS RESPOND TO INTERVIEW QUESTIONS?

An interview is not a very fruitful information-gathering technique if interviewees cannot answer questions. In view of the limited verbal skills of mentally retarded persons, we decided that the first challenge in interviewing the retarded is to obtain a response of some kind. Thus we measured and analyzed the responsiveness of interviewees, their ability to meet the demands of a particular question with an answer suited to the form of the question (e.g., "yes" or "no" to a yes-no question) and relevant to the content of the question (e.g., about work if the question is about work). We were guided by the following questions:

1. What can be expected of mentally retarded persons in interviews, and what types of respondent behaviors constitute problems in interviewing the retarded?
2. To what extent is responsiveness to an interview a predictable individual behavior, and what factors are related to it?

3. How can responsiveness be optimized?

To convey what we have discovered about responsiveness, we will take up the following topics in turn, drawing from the various studies to discuss each topic: specific types of responses, the stability of responsiveness, the relationship of responsiveness to IQ and other client characteristics, the relationship of responsiveness to question type, and the optimization of responsiveness.

Specific Types of Responses

In most of our analysis, we dealt with a measure of the percentage of questions asked of a respondent which were answered at least minimally appropriately; that is, which met the formal and substantive demands of the question. However, as indicated in Chapter 3, we also measured various types of inappropriate or uncodable responses, and were interested in the extent to which each might pose problems in an interview.

Table 5.1 presents the mean percentages of responses which fell in each of the nine responsiveness codes in the two institutional samples. In each case they are based on responses to one of the two parallel interview schedules used in these studies (82 questions for the children's study and 142 questions for the adult study). Profoundly retarded adults were excluded because the vast majority of them failed the screening interview and were thereby defined as totally unresponsive.

Their exclusion makes the child and adult samples comparable in IQ ($M=42.08$ for children, 39.76 for adults). The means in Table 5.1 cannot always be taken as representative of typical performance because some distributions were unusual (for example, some clients never gave a response while some others never failed to respond in some way). However, they are suggestive of the frequency with which various types of interview behavior can be expected to occur.

Table 5.1: Mean Percents of Responses Falling in Each Category in Institution Samples

Responsiveness Code	Adult Institution (N=42) *	Child Institution (N=52)
1. No response	10.20	6.55
2. Unintelligible	4.71	4.91
3. Irrelevant	3.48	5.11
4. Don't know	.28	2.85
5. Inadequate	3.73	7.41
6. Request for more information	1.01	3.32
7. Refusal to answer	.01	.33
8. Minimally appropriate	63.63	57.94
9. Expanded appropriate	12.99	11.58
	100.00%	100.00%

*Excludes those with IQs below 20, who were generally totally unresponsive, to make groups more comparable in IQ.

We can interpret these figures either of two ways. We can say, for example, that the typical child interviewee was able to give a minimally appropriate answer to 63.63% of the questions on the interview schedule. Alternatively, we can say that of all the behaviors elicited from the sample by interviewers' questions, 63.63% of them were minimally appropriate answers. Clearly in both samples we obtained minimally appropriate or expanded answers in a majority of cases, more often minimal than elaborated.

What of the inappropriate types of respondent behavior? Generally, a high percentage of inappropriate responses took the form of no response at all. Unintelligible responses were also a problem, unintelligibility, of course, being a function of the interviewer's or listener's capabilities as well as of the interviewee's communication skills. It is highly likely that a person who knew the respondents might well have been able to decipher many responses that our interviewers had to judge unintelligible, but in most surveys interviewer and interviewee are not familiar with one another. Irrelevant responses, that is, answers that appeared unrelated to the substantive demands of the question, were a frequent problem as well, often because interviewees did not understand a particular question but in some cases because interviewees had a general lack of comprehension of the interview but chose to speak. Somewhat more common still were inadequate responses, which were judged to be relevant to the question but too vague or ambivalent to be coded into content categories. At times respondents indicated that they did not know the answer, but this was fairly rare. Our coding system was based on an assumption, usually but not always valid, that respondents should have known how

to answer the questions we asked. Similarly, interviewees occasionally requested more information about the question before answering, usually by saying "huh?" to elicit a repetition of the question, but this did not happen often either. By far the rarest event was a refusal to answer the question, even though participants had been told at the start that they did not have to answer any question they did not want to answer.

The relative frequencies of the nine types of responsiveness codes were much the same in the two institutional samples as well as in other samples. We would have to conclude that interviewers of the retarded can expect to encounter a wide range of inappropriate responses, but that appropriate responses can generally be expected a majority of the time.

The Stability of Responsiveness Scores

In both of the institution samples, similar interview schedules were administered a week apart, allowing us to examine the extent to which responsiveness is a consistent individual behavior. In the children's institution sample, Form A consisted of 82 questions and Form B of 65 questions, with 31 questions repeated in exactly the same manner on both administrations and the remainder of the questions typically representing alternative ways of seeking the same information. A total responsiveness score on each form was calculated for each subject (the number of questions answered appropriately divided by number of questions asked). When responsiveness to Form A and responsiveness to Form B were correlated, the result was a highly significant correlation of .96, indicating that the extent to which an interviewee was responsive to one form could be predicted almost per-

fectly from his or her responsiveness to the other form. Although the correlations were lower, the frequencies of specific types of response on Form A and B were also significantly related, with the exception of the category "Request for More Information." These correlations were .54 for no response, .74 for unintelligible, .39 for irrelevant, .30 for don't know, .33 for inadequate, .20 for request for more information, .92 for minimally appropriate, and .54 for expanded appropriate responses. Thus, the extent to which various less-than-appropriate and appropriate responses occurred was somewhat predictable from time to time.

In the adult institution sample, this finding of stability was confirmed. Indeed, the correlation between total responsiveness scores for the two forms (Form A, 142 items, Form B, 131 items) was identical (.96).

As another way of examining the stability of responsiveness scores, we analyzed the extent to which the group's level of responsiveness

changed from first to second administration. In the children's institution sample, this analysis focuses on 31 questions repeated exactly the same on both administrations. A Latin square design was used to examine the administration effect in the context of interviewer effects (see Table 5.2). Overall, there was not a significant difference between first and second administration responsiveness, although scores did increase somewhat from first to second administration (from 71.8% to 74.6% overall). There was also no overall interviewer effect, that is, neither interviewer elicited higher rates of responsiveness from the sample. There was however, an interaction between interviewer and administration, $F(1, 49) = 9.92, p = .005$. Interviewer Two, for some reason, was associated with a greater increase in responsiveness from first to second administration than was Interviewer One. It is possible that the 26 subjects whom she interviewed on second administration were, for reasons we cannot explain, more susceptible to practice effects than were the 26 subjects who had Inter-

Table 5.2: Responsiveness of Children As a Function of Administration and Interviewer*

	Interviewer	
	One	Two
Administration One	72.20	71.46
Administration Two	73.92	75.28

*For an N of 52 in severe to mild ranges. One half of the subjects had Interviewer One first and half had Interviewer Two first, and then the other interviewer conducted the second administration.

viewer One on second administration. Or it is possible that Interviewer Two became more relaxed, warm, or stimulating in manner in the second round of interviews.

A similar analysis for the adult institution sample (see Table 5.3) revealed no significant differences in responsiveness associated with administration, interviewer, or the interaction of the two. Again, however, responsiveness did increase slightly from first to second administration, and this effect was more noticeable for Interviewer Two.

To put these lines of evidence together, then, we would conclude that responsiveness to interview questions is a stable individual behavior. Although some respondents may increase in their ability to respond to questions as a result of practice in a prior interview, dramatic changes are not to be expected. Moreover, an individual's ranking in the group as more or less responsive appears to be highly consistent from interview to interview.

Responsiveness and IQ

It is quite clear from the literature on language development among the mentally retarded (see chapter 2) that in the lower ranges of the IQ continuum, more delay in language and speech is to be expected, and more totally nonverbal individuals are encountered. A critical question in our study of the feasibility of interviewing mentally retarded persons centered on clarifying what can be expected at different levels of retardation. At what point in the IQ continuum does obtaining answers to questions become infeasible?

We set out with low expectations for persons in the profound range of retardation, but nonetheless attempted to interview persons with IQs below 20 in the study of institutionalized adults. It was here that our simple screening procedure, described in Chapter 4, was heavily used to avoid the unpleasant experience of holding an interview when it was apparent that the subject could not participate. As noted, the rule was to accept any verbal or nonverbal response, whether intel-

Table 5.3: Responsiveness of Adults as a Function of Administration and Interviewer*

	Interviewer	
	One	Two
Administration One	66.15%	63.37%
Administration Two	67.82%	68.23%

*For 42 interviewees in severe to mild range. One half of the subjects had Interviewer One first and half had Interviewer Two first, and then the other interviewer conducted the second administration.

ligible or not, as sufficient grounds for attempting the full interview. Thus, the screening criteria were designed to give anyone who showed the least sign of ability an opportunity to be interviewed.

Of the 16 profoundly retarded adults who were approached, 13 or (81%) failed the screening and were judged totally incapable of participating in an interview. The responsiveness of the three who did pass the screening was deficient. Two of them were able to respond adequately to slightly over half of the questions, primarily yes-no questions. The third subject failed to give an answer two-thirds of the time and otherwise gave a high proportion of irrelevant answers.

On the whole, then, our expectations about the infeasibility of interviewing profoundly retarded persons were borne out. While three such persons were able to respond to some extent, we are not confident that their answers were meaningful and they were the exception rather than the rule. Any attempt to seek information from such persons should rely on other methods of data gathering; systematic behavioral observation of circumstances, activity patterns and expressed affect, and, where feasible, use of manual and other nonverbal communication systems.

All three of the major samples in the study (institution children, institution adults, and community children) included persons in the severe to mild ranges of retardation. Thus, we were able to examine in each sample the relationship between IQ and ability to respond to interview questions for persons with IQs ranging from 20 to 68.

In the institution children's sample, the ability to respond was quite clearly a function of IQ. Table 5.4 represents the mean percentages of

responses falling in each of the responsiveness codes (with the exception of refusals to respond, which were so rare that they are not reported). It can be noted that instances of no response, unintelligible responses, and irrelevant responses decreased as IQ increased, while the percentage of appropriate responses, both minimal and expanded, increased. In several respects, severely retarded subjects, as a group, were quite different from moderately, and mildly retarded ones. Most notably, they were much more likely to fail to respond and much less likely to provide appropriate responses. Overall, the percentages of appropriate response to Form A for severely, moderately and mildly retarded groups were 52.9%, 89.6%, and 93.3% respectively, a highly significant IQ group effect, $F(2.49) = 22.51$, $p = .0001$.

However, it is important to note that variability was also greater in the severely retarded group ($SD = 38.8$) than in the moderately retarded ($SD = 9.1$) and mildly retarded ($SD = 3.7$) groups. The responsiveness of individuals in the severe retardation range was, in short, relatively unpredictable. Four severely retarded children failed the screening, while the most responsive gave appropriate answers 96.3% of the time, and the rest spanned the range in between. By contrast, Form A responsiveness scores in the moderately retarded group ranged from 72.2% to 100%, and in the mildly retarded group from 83.7% to 97.5%, suggesting that all children in these IQ ranges were able to answer most of the questions they were asked. What this suggests is that one cannot make firm predictions about whether or not interviewing is feasible with severely retarded persons. The only way to tell is to attempt an interview and see how individuals respond.

Table 5.4: Responsiveness Codes by IQ Group:
Mean Percents of Answers on Form A

	Severe (N=20)	Moderate (N=16)	Mild (N=16)
No response	25.8	0.9	0.0
Unintelligible	8.9	4.0	.2
Irrelevant	8.0	1.1	.2
Don't know	0.1	0.2	.6
Inadequate	3.3	3.7	4.3
Request more information	1.1	0.6	1.3
Minimally appropriate	50.0	72.7	71.6
Expanded appropriate	2.9	15.9	21.7
	100.1	100.1	99.9

For this sample the overall correlation between responsiveness of Form A and IQ was .62, and between responsiveness of Form B and IQ .76, again indicating a strong relationship between IQ and responsiveness. At the same time, there was a negative relationship between IQ and the percentage of "no response" codes ($r = .59$ for Form A, $-.52$ for Form B) as well as between IQ and the percentage of all other inappropriate responses combined ($r = .36$ for Form A, $-.44$ for Form B).

In the adult institutional sample, much the same pattern of findings emerged. Overall, of Form A, severely retarded persons averaged 58.8% responsiveness, and were less responsive than moderately (78.0%) and mildly retarded (73.5%) groups, $F(2, 39) = 3.63$,

$p = .04$. On Form B, the same pattern emerged (severe, 56.3%, moderate 76.6%, and mild 75.5%, $F(2, 39) = 4.46$, $p = .02$. In this sample, one severely retarded person failed the initial screening. In addition there was one mildly retarded subject whose emotional disturbance seriously interfered with participation in the interviews and whose scores lowered the mean for the mildly-retarded group and increased the extent of variability. This woman was not able to respond to 90.3% of the questions on Form A and 85.3% of those on Form B. Her exclusion from the analysis increased the mean of Form A for the mildly retarded group from 73.5% to 81.1% and decreased the standard deviation from 25.5% to 10.1%. As in the children's institution study, variability in the severe retardation

range was high (25.4% compared to 11.7% in the moderate range and 10.1% in the mild range with the outlier excluded.)

In correlational analysis, the relationship between IQ and percentage of appropriate responses on Form A was a significant but relatively low .35 with the one outlier in the mildly retarded group included (and .51 with her score excluded). As expected, given the high relationship between responsiveness scores for the two forms, results for Form B were similar (.42 with her included, .57 with her excluded). Again, then, the study of institutionalized adults indicated that the ability to respond to questions is a function of IQ, with responsiveness dropping off markedly in the severe range of retardation, but with severely retarded persons differing widely from each other, some being incapable of participating at all and others responding much like moderately and mildly retarded persons.

Our third opportunity to examine the relationship between IQ and responsiveness was in the children's community sample. The interview schedule used in this study was, in our opinion, easier than those used previously because we had by then analyzed responses to the questions used in the institution samples and attempted to refine techniques to optimize responsiveness. In this sample, responsiveness was indeed higher. Severely retarded persons did relatively well, averaging 78.8% responsiveness, compared to 82.1% in the moderately retarded group and 92.7% in the mildly retarded group, $F(2, 54) = 5.93, p = .005$. In this case, both the severely and moderately retarded groups were found to be significantly less responsive than the mildly retarded group. The overall correlation between responsiveness and IQ was .41. Moreover, in this

sample, the extent of variability among severely retarded persons was not higher than that among the two higher IQ groups ($SD = 12.8$ for severe, 18.4 for moderate, and 5.0 for mild). In our opinion, by reducing the number of open-ended and multiple choice questions and otherwise working to simplify the types of questions that were asked, we may have succeeded in making the task simpler for the community sample and reduced the extent to which severely retarded persons appear to be deficient compared to their higher IQ peers. For this reason, these figures do not provide a basis for concluding that severely retarded persons in the community are more verbally skilled than those in an institution.

However, we were able to compare the three samples more directly by examining their responsiveness to 22 questions asked in exactly the same way in all three samples. These questions included seven verbal yes-no questions, nine yes-no questions about chores accompanied by pictures, two either-or questions, one multiple choice question, and three open-ended questions.

Table 5.5 presents the average percentage of questions answered appropriately in each of the three samples at each of the three different levels of mental retardation. As the table suggests, the data for the different samples were somewhat different, and this was reflected in a significant interaction effect in a 3×3 analysis of variance between sample and level of retardation, $F(4, 140) = 4.60, p = .002$. This finding required that we look more carefully at IQ group differences within each sample. In the children's institution sample, the IQ group effect was significant $F(2, 49) = 14.48, p < .001$, and t-tests indicated that severely retarded persons were less responsive on the average than both moderately and mildly retarded

Table 5.5: Responsiveness by IQ Group for 22 Questions
Used in Three Samples

	Level of Retardation		
	Severe	Moderate	Mild
Institution children	55.17	94.11	93.35
Community children	84.59	85.56	95.68
Institution adults	71.59	90.91	78.41

persons. In the children's community sample, IQ again mattered, $F(2, 52) = 3.75$, $p = .03$. Here, however the severely retarded group was significantly less responsive than the mildly retarded group and the moderately retarded group was almost so ($p = .065$). In the adult institution group, by contrast, IQ group differences overall were not significant, $F(2, 39) = 2.54$, $p = .09$, but follow-up t-tests did indicate that severely retarded clients were significantly less responsive than moderately retarded clients ($t = 2.87$, $p = .009$). In this sample, primarily because of the almost total unresponsiveness of one emotionally disturbed, mildly retarded woman, the mildly retarded group did not have the highest average responsiveness. Thus IQ group differences were weakest in this group and strongest in the institution children's sample.

The one consistent conclusion to be reached is that severely retarded interviewees generally can be counted on to be less responsive than their higher IQ peers, but the specific magnitude and locus of IQ group difference appear to vary from sample to sample. Over all samples, of course, the effect of IQ group

was significant, $F(2, 146) = 15.55$, $p = .001$, with severely retarded persons averaging 62.99% responsiveness, as compared to 72.3% for the moderately retarded and 77.72% for the mildly retarded.

What of the comparison between samples? Overall means for institution children, community children, and institution adults were 77.45%, 88.92%, and 79.33%, respectively, and differences between samples approached statistical significance, $F(2, 140) = 2.63$, $p = .08$. However, since mean IQs for the three samples were different, analysis of covariance was performed to examine differences between samples with IQ controlled. In this analysis, the samples were not significantly different in their responsiveness, $F(2, 145) = 1.16$, $p = .32$. In other words, the adult sample had no clear advantage over samples of children, and, more importantly, the apparent advantage of community children over institutionalized persons vanished when IQ was controlled. Thus, contrary to some evidence in the literature, institutionalized persons were not found to be notably deficient compared to community residents in communication skills required to answer questions.

Relationships of Responsiveness to Other Personal Characteristics

Although we consistently found that responsiveness was related to IQ, we found no consistent evidence that responsiveness was a function of other client characteristics. It is most useful to focus on the 22 questions analyzed above which were repeated in all three samples. Correlations between sex and percent of the 22 questions answered adequately in the three samples ranged from $-.02$ to $.11$, suggesting no relationship between sex and responsiveness. Similarly, correlations between age and responsiveness were weak ($-.09$ to $.12$). Of course, because the age range in the two children's samples was restricted to 12-16, there was not much opportunity for a relationship to manifest itself in these groups. However, the correlation of $-.09$ in the adult samples where ages ranged more widely ($SD = 4.41$) was not significant either. As we noted above, there were no recognizable differences between institutionalized and noninstitutionalized groups either, contrary to some evidence in the literature of language deficits among the institutionalized. In short, our search for other subject characteristics predictive of the ability to respond in an interview situation was not productive. However, we did find relationships between responsiveness and the type of question asked, an important set of findings to which we turn next.

Responsiveness and Question Type

To illustrate in some detail the technique used to analyze differences in responsiveness as a function of question type, we will first examine the relationships existing within the institutionalized children's sample and then explore the extent to which it was similar in the other sample.

RESPONSIVENESS AMONG INSTITUTION CHILDREN AS A FUNCTION OF QUESTION TYPE

While most of the children's institution interview consisted of simple yes-no questions, there were other types of questions used which permitted a comparison of severely, moderately, and mildly retarded children's ability to respond appropriately to questions as a function of question format. Six scales were formed, each measuring the percentage of questions of a given format that were answered with at least a minimally appropriate response. The subsets of items used to calculate these scores were as follows:

1. *Factual Yes-No Questions.* Ten questions about activities (Do you: ... read books, magazines, or newspapers; go out to the movies; go out to eat; go to church; go to stores; watch TV; listen to the radio or record player; play any sports; go to school or take classes; do any arts and crafts like ceramics or painting?)
2. *Subjective Yes-No Questions.* Eight yes-no questions, consisting of four pairs on the same topics, included as part of an attempt to study acquiescent response set (Are you usually happy? Are you usually sad? Do people here treat you nice? Do people here treat you mean? Is this place where you live big enough? Is this place where you live too little? Are you usually by yourself? Are you usually with other people?)
3. *Either-Or Questions.* Eight either-or questions covering the same topics as in scale 2 above; four pairs with different orderings of terms (e.g., Are you usually happy or sad? Are you usually sad or happy?).

4. *Verbal Multiple-Choice Questions.* Three questions, each supplying four answers from which to select (How often does anybody in your family come to see you? A lot, sometimes, not much, or never? How many friends do you have? A lot, some, not many, or none?)
5. *Pictorial Multiple-Choice Questions.* Three questions, each of which asked the respondent to select a happy face, slightly happy face, slightly unhappy face, or unhappy face as descriptive of his or her feelings (Which picture shows how you feel about living here? Which picture shows how you like the food here? Which picture shows how you like the people here?)
6. *Open-Ended Questions.* Five open-ended items (If you had one wish, what would you wish for? Who helps you the most when you need help? What are they teaching you there (at school)? Where do you get your money? What do you and your friends usually do together?)

With the exception of the yes-no items which were sampled, the rest of these scales included all items on the two forms of the questionnaire fitting each format. Our hypothesis was that yes-no questions and pictorial multiple-choice questions would be associated with the highest responsiveness rates, and that the other questions with closed formats (either-or and verbal multiple-choice) would elicit higher rates of response than would open-ended questions which require that respondents frame the response.

Figure 5.1 shows the mean percentages of appropriate responses on each type of question for each IQ group. This analysis by question subset indicated, first of all, that there were significant differences among responsiveness rates in the total sample as a function of question type ($F(5, 45) = 23.90, p = .0001$). For the 50 subjects who had subscale scores for each of the six formats (two visually impaired persons could not complete the pictorial multiple-choice section), this analysis of variance by question type was followed up with t -tests for dependent measures to locate specific differences among question types (see Table 5.6). Factual yes-no questions and pictorial multiple-choice questions were found to be significantly easier to answer than all other types of questions and were not significantly different from one another. The subjective yes-no questions were significantly more difficult than factual yes-no questions ($t(49) = 3.41, p = .001$) and bordered on being more difficult than pictorial multiple choice questions as well ($t(49) = 1.95, p = .057$). Either-or questions ranked next in difficulty, being associated with significantly higher responsiveness rates than verbal multiple-choice and open-ended questions. Correlations among the six subscales (see Table 5.7) indicated that, generally, children's levels of responsiveness on one type of question were predictive of their responsiveness on other types. All of these correlations were significant well beyond the .05 level. However, correlations among simply structured questions were generally of a larger magnitude than were correlations involving open-ended questions or the four-choice verbal multiple-choice questions.

The analysis points out the advantages of using yes-no and pictorial multiple choice questions when the goal

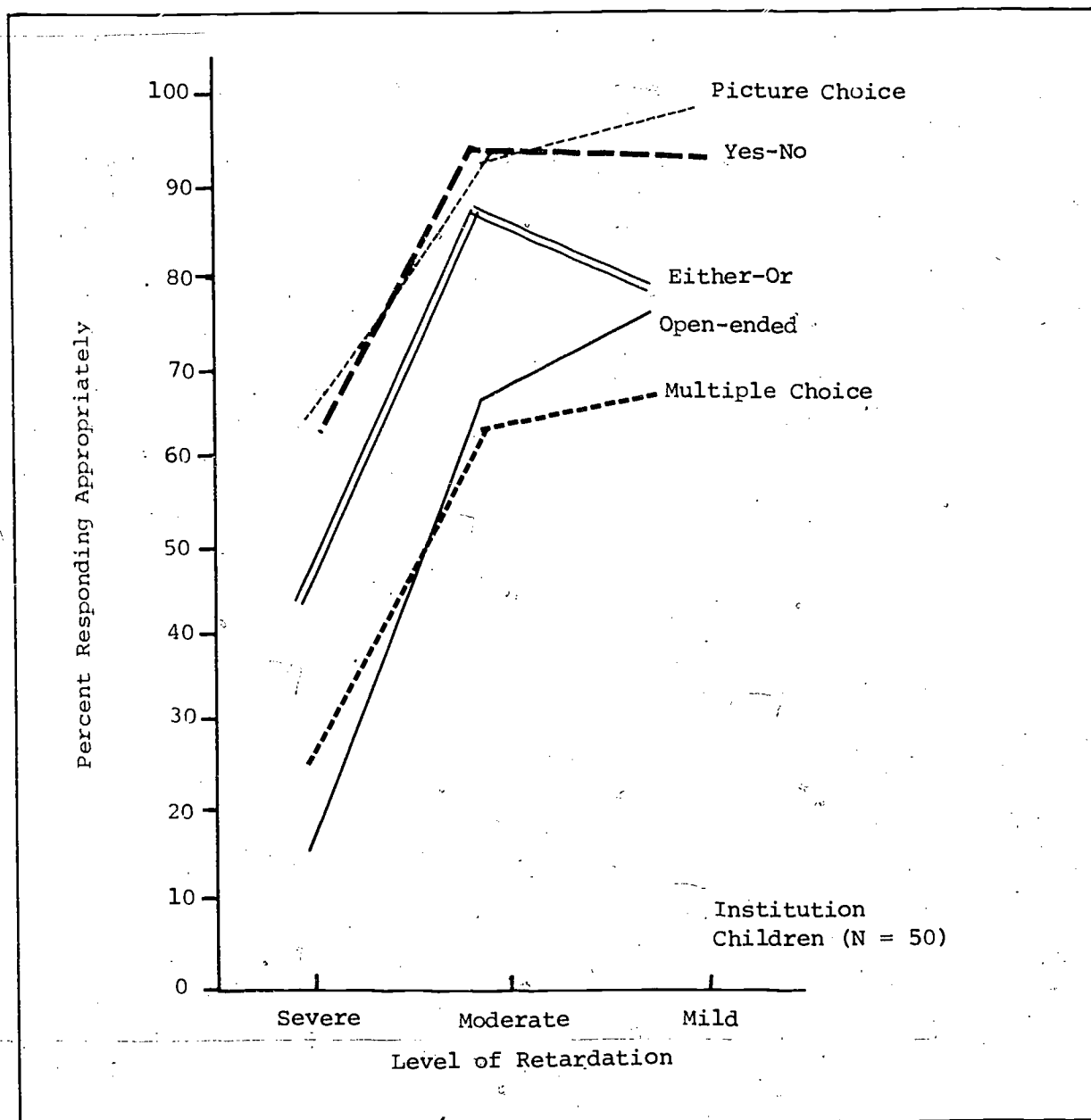


Figure 5-1 Responsiveness to Various Question Types as a Function of Level of Retardation.

is to obtain responses from as many mentally retarded interviewees as possible. As the difference between factual and subjective yes-no questions indicates, the content of the question has some bearing on the likelihood of obtaining a response,

but is relatively minor compared to the effect of question format. Subjective yes-no questions, and perhaps the subjective either-or questions used in this study as well, often forced respondents to make a black-and-white choice. Where responsive-

Table 5.6: T-Values for Differences between Appropriateness of Response Scores on Six Subsets of Items (N=50)

	Pict.	Fact Y-N	Subj. Y-N	E-Or	Verbal	Open
Pictorial multiple-choice		.35	1.95	4.27*	6.14*	6.37*
Factual yes-no			3.41*	4.28*	6.43*	6.30*
Subjective yes-no				2.38*	5.20*	4.97*
Either-cr					4.17*	4.15*
Verbal multiple-choice						.25

*Significant at the .05 level.

ness fell off, it was often because residents would request clarification of the question or give vacillating responses which could not be coded one way or the other. While pictorial multiple-choice questions were easy to answer, undoubtedly because they required only pointing and not necessarily understanding, verbal multiple-choice questions were, surprisingly, as difficult as open-ended questions. While this may have been

Table 5.7: Correlations Among Appropriateness of Response Scores on Six Types of Questions

	Pict.	Fact Y-N	Subj. Y-N	E-Or	Verbal	Open
Pictorial multiple-choice		.77*	.72*	.75*	.48*	.57*
Factual yes-no			.93*	.78*	.52*	.53*
Subjective yes-no				.76*	.49*	.49*
Either-or					.62*	.67*
Verbal multiple-choice						.54*

*Significant at the .05 level.

partially due to the topics and quantitative response options for the questions used in the study, it is quite likely that four options are too many for many mentally retarded persons to consider at once. Certainly the two-choice either-or questions were much easier by comparison, despite the fact that they were subjective and relatively abstract.

Analysis of IQ differences in responsiveness to each of the six types of questions was conducted through one-way analysis of variance with follow-up *t* tests. Separate variance estimates were used due to higher variances in the severely retarded group than in the moderately and mildly retarded groups. On all six sets of questions, IQ group differences, as indicated by *F* statistics, were significant beyond the .005 level. Moreover, follow-up *t*-tests indicated that for all six question formats, the severely retarded group was less responsive than the moderately and mildly retarded groups and the latter two groups did not differ significantly from each other. The only difference between the moderately and mildly retarded subjects approaching significance was on the either-or questions, where, contrary to expectation, the moderately retarded were somewhat more responsive than the mildly retarded (*t* (28) = 1.83, *p* = .078). All differences between the severely retarded and the other two groups were significant beyond the .01 level.

Thus, the IQ group differences in total responsiveness for this sample, which pointed to deficiencies among the severely retarded, do not appear to be a function of the fact that most of the questions asked in the study were in yes-no form. The severely retarded appeared to be less capable of giving appropriate responses to all types of questions asked. Of course, their low responsiveness rates on verbal multiple choice and open-ended questions, and

to a lesser extent either-or questions, raise serious questions about whether those formats should be used at all with severely retarded children, for in each case, less than half of the severely retarded group studied was able to respond appropriately. Moreover, it was on these types of questions that the difference between the severely retarded and the high IQ groups appeared to be most pronounced.

RESPONSIVENESS AND QUESTION TYPE IN OTHER SAMPLES

Table 5.8 presents the mean responsiveness percentages for the same categories of questions (although not exactly the same questions) in the children's institution, adult institution, and child community samples. One can readily see that the ordering of the question formats is much the same in all three samples. Among institutionalized adults, factual yes-no questions about activities were slightly easier than yes-no questions of a subjective nature, and both types of yes-no questions were significantly easier than verbal either-or questions, which in turn elicited higher responsiveness scores than either verbal multiple-choice or open-ended questions. Thus these findings paralleled almost exactly those for the institutionalized children, either-or questions being more difficult than yes-no questions but easier than the challenging open-ended and verbal multiple choice questions.

The pattern of differences in the community children's sample was somewhat different. Pictorial choice and factual yes-no questions about activities were significantly easier than all other types of questions. Subjective yes-no questions, while still falling next in order, were significantly more difficult than these two formats but yielded higher responsiveness scores than did the remaining formats. Verbal either-or questions, rather than being easier than verbal multiple-

Table 5.8: Mean Percentages of Appropriate Response to Various Types of Questions in Three Samples

Question format	Sample		
	Institution Children (N=50)	Institution Adults (N=42) *	Community Children (N=57)
Pictorial choice	83.3	(not used)	96.1
Verbal yes-no (factual)	82.2	84.8	94.0
Verbal yes-no (subjective)	76.5	81.7	81.7
Verbal either-or	68.7	65.7	72.5
Verbal multiple-choice	49.7	51.9	72.9
Open-ended	51.0	50.6	69.9

*Excludes profoundly retarded group.

choice and open-ended questions, proved to be as difficult. In this sample, responsiveness was generally high, and there was not the marked drop in responsiveness associated with verbal multiple-choice and open-ended questions as there had been in the institutional samples.

Still, the ordering of question formats must be considered highly stable from sample to sample. One can generally expect to obtain answers from a large proportion of a sample of retarded persons if one asks picture choice and verbal yes-no questions. Not coincidentally, both types of questions can be answered nonverbally (by pointing in the first case, by nodding or shaking the head in the second).

The use of verbal either-or questions involves some sacrifice of responsiveness, but still can be expected to yield appropriate answers from roughly two-thirds of the sample. To our surprise, verbal multiple-choice questions were consistently as difficult to answer as the more obviously difficult open-ended questions. In our interviews, these questions typically required a choice among four, or sometimes three, alternatives (e.g., "never," "not much," "sometimes," or "a lot" in answer to a question about the extent of involvement in some activity). Apparently, either because a quantitative concept was involved or because subjects were unable to process so many options, these questions proved as difficult as questions that

required respondents to generate responses independently, with benefit of structure.

It is difficult to fully convey the problems we encountered with open-ended questions without providing concrete examples. Samples 5.9 and 5.10 provide verbatim answers given by virtually all members of the adult sample who could respond verbally in some manner. (We excluded "don't know" and "huh?" responses as well as a few repetitive answers.) The tables illustrate how the ability to answer such questions improves as IQ increases, but they also illustrate how few responses to open-ended questions were interpretable. Many adults misunderstood the questions. For example, several, instead of indicating what people did that bothered them, named people who bother them, and some seemed to think that they should discuss things they did that are nice or bothersome, or simply things that they like or dislike. Where respondents understood the questions, their answers did provide some clues to how they view the social world around them (e.g., note the comments about noise as a source of irritation), but the insights to be gained are limited. Particularly for questions for which categories were constructed after client and significant other responses had been reviewed, it proved extremely difficult to sort answers into meaningful categories. Moreover, as the tables suggest, once one omits the irrelevant or inadequate (vague) responses, very few codable answers are left to categorize in the first place. Again, while insightful answers sometimes emerge from more verbal respondents, the return on investment, if we may borrow a phrase from business, is generally prohibitively low for open-ended questions.

What of the differences among IQ groups in their ability to manage different question formats? In the adult institution samples, differences between IQ groups were not significant for subjective yes-no questions, factual yes-no questions about activities, and multiple-choice questions, although trends pointed toward higher responsiveness as IQ increased. On both either-or questions and verbal multiple-choice questions severely retarded persons were significantly less responsive than moderately retarded persons. (These analyses included the mildly retarded subject who was generally unresponsive and who lowered the mean for mildly retarded subjects.) On open-ended questions, the severely retarded group was even more seriously deficient, averaging 30.3% responsiveness compared to 62.7% for moderately retarded persons and 63.9% for mildly retarded persons. Thus, in this sample, while the severely retarded group invariably had the lowest mean responsiveness, their most notable deficiency was on open-ended questions.

Among the community children, a somewhat similar pattern emerged. While there were no significant differences on yes-no questions about activities, picture choice questions, and subjective yes-no questions, severely retarded subjects were significantly less responsive than both higher IQ groups to verbal either-or questions. On verbal multiple-choice questions, they were significantly less responsive than were mildly retarded subjects. And finally, once again, the most striking IQ-group differences emerged for open-ended questions. Here, severely retarded respondents, with a mean responsiveness score of 51.1%, were significantly less responsive than the moderately retarded (68.9%), who in turn were less responsive

Sample 5.9

"What Nice Things Do people Here Do for You? What Else?"

Selected Responses of Institutionalized Adults*

Profound

It's from playing games on the record hop

Yeah

Severe

Nice

Like to Work

No way

Work

Mrs. ___, she's my houseparent, M. S. ___, etc. (list names)

Three

Mary

I don't do that. /Sometimes you talk to people/ and if they go out,
that all. And sometimes we go skating/Sometimes we dance:
together.

We talk, that's all.

I used to play around and ride my bicycle.

Nice

Moderate

Play with me/play games/play hide and seek/farmer in the dell and
hopscotch and play jacks/answer telephone

They give me books and they give me necklace/give me home on weekends/
well, buy me things/I don't think of no more.

I do myself, help me out.

Get me up in the morning and feed me my breakfast/They help me get
on the pot and off the bathroom/They feed me sometimes/They help
me go to school/They're so nice to me/They treat me nice here.
They let me go to the canteen when I want/Let me buy something when
I want to/Come to cottage and drink a coke and stuff, take a bath
and get ready for supper/Just go swimming whenever we want/Sweep
during P.M.

Let us go shopping/after movies/bowling

Oh, not much

Let me buff the floor/Mop

They help you out/Show you how to do things/(unintelligible) and make
the bed

They work here and clean up, mop and clean and kiss.

Give us treats

I help in the kitchen/I take my plate up

Nice peoples. They been good, quiet.

Usually they're so tied up. But when they have time, I usually go
someplace with them/I made my bed this morning

Mild

Work/Stay work/wash face, hair/shave/try hair spray

Help me out, Help me out when I'm sick

They do, too--buy me things, everything I need/They let me do things I
want to/They let me go places

Try to help me get a job/Help every people

They help me walk/They go to the canteen for me

They tell me when it's time to eat and go to class and sometimes we

just sit down and tell each other jokes

Wash your clothes/Talk to you. Ask you questions, talk about things
you want to talk about. Look at things you want 'em to/Sew
your clothes, talk to you, put things up for you, pick things
up for you/Remind you to do stuff

Well, I help make beds and strip beds and dress the girls. And if
I'm nice, I stay in the back and help ladies strip beds. If I
stay back there, they don't have to yell at me and they let me
have a cup of coffee for breakfast and they let me wash dishes/
When they're up in front, I'm usually washin' and dryin' clothes,
when I lived in ____... (goes on to talk about punishments).

*Slashes mark points at which the probe, "What else?" was provided.

Sample 5.10

"What Do People Do Here That Bothers You? What Else?

Selected Responses of Institutionalized Adults*

Profound

Yeah

It's from bothers you, bother, bother

Severe

Fight

Did you find my screw? (for his glasses)

Nobody bothers me

Roy bothers me. I buy him a coke, he's good (rambling on)

Kevin, George, James

Nobody

Alice does sometimes

They do. I slap people/There's something they do and I hit 'em.

Work

Calls me _____ and I don't like that name.

Moderate

Nothin'

Oh, some bother me. I get along with some, not all.

Make noise/They yell/They keep it

I hate to be picked on/noise/that's it

They pass me around

Fuss

I don't like going around doing stuff/Everything bothers me/Pickin up
trash.

Knock in the head/Knock down; soup when we have soup

Gets you in trouble/and play

Nothing, noise don't hurt me.

Chews tobacco and stuff/smoking/pestering when I get off from laundry/

I can rest. I'm tired when I get home from work but they disrupt
me/stealing, I can't stand that/that's it

Mild

Saturday.

Nothing

Bad man

Call names, cuss; they cuss and talk mean/(unintelligible)/stare at
you/they kick you every day

Talk a lot/when I try to sleep, they wake me up. That bothers me more
than anything in the world.

I don't like real loud talk. I don't like fussin' or fightin'/
hittin' people/one thing they have on my medical folder, that
whenever I was born something happened to my mind. They say
they can't get my attention as fast as they want to. I really
want to go home for good. I believe I'm old enough.

They hit on me and they call me names and they bite me and they get
me in trouble/nearly every morning, every afternoon and every
night, I have to stand this screamin' and yellin' that goes on
every afternoon, morning, and night, and I get this headache
that won't go away. I called the housemother and asked her to
ask the nurse to give me a couple of aspirins. But the only
person in this cottage who gets aspirin is ____.

Sometimes they might ask questions I can't answer, or a few other things. I don't get bothered real easy-nothing really.

*Slashes mark points at which the probe, "What Else"," was provided.

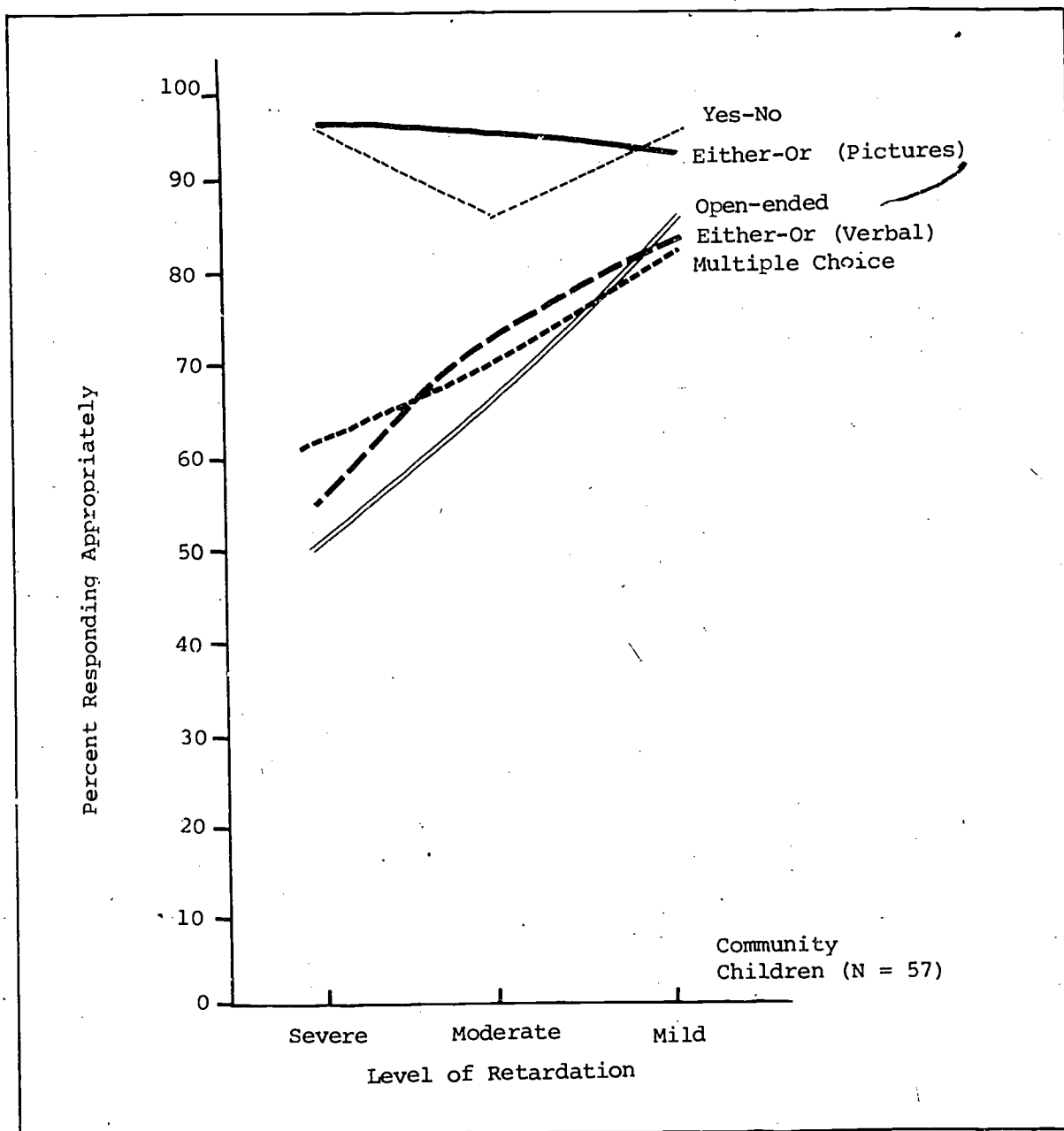


Figure 5.2 Responsiveness to Various Question Types as a Function of Level of Retardation

than the mildly retarded (88.5%). These findings are strikingly portrayed in Figure 5.2.

The pattern of results, then, suggests that while severely retarded persons can generally be expected to have the most difficulty responding

to questions, their handicap is most evident on the questions which can be considered the most cognitively demanding, particularly open-ended questions which put the burden for generating an answer on the interviewee. These findings, collectively,

have clear implications for survey design where the goal is to optimize responsiveness. We will consider another means of optimizing responsiveness next.

Effects on Responsiveness of Asking Questions Twice

Certainly one possible reason for low responsiveness is the interviewee's failure to process the question. Throughout our research, we used responsiveness to the first asking of the question as the major measure of responsiveness. However, we also used the strategy of asking the question twice if the interviewee had not been able to answer it

on first asking. Moreover, in the adult institutional sample, we explicitly tested the value of repeating the question as a way of increasing responsiveness.

The analysis focused on a total of 131 different questions from both forms which were asked of all interviewees. The measures compared were the percentage of the sample responding with at least a minimally appropriate response after both first and second asking. These two figures were compared through *t* tests for correlated means for each of five sets of questions: 82 yes-no questions, 20 open-ended questions, 11 multiple-choice questions, 13 picture choice ques-

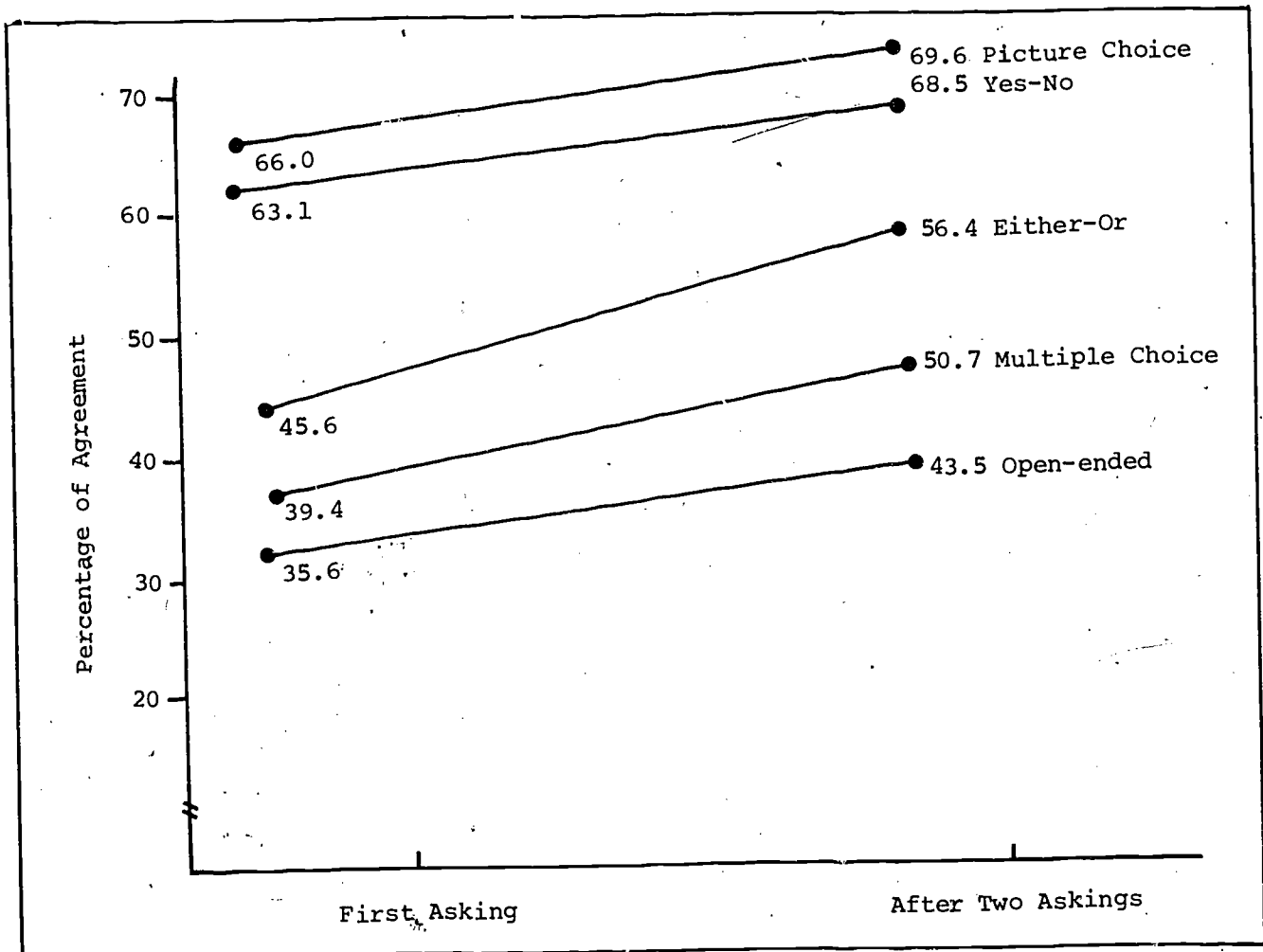


Figure 5.3 Percent of the Sample Responsive to Various Question Formats After One Asking and After Two Askings

tions, and 5 either-or questions. Figure 5.3 shows the mean responsiveness figures for each of these five sets at first asking and for first and second asking combined.

The figure clearly indicates that for each category of question there was at least some gain in sample responsiveness figures associated with giving interviewees a second chance to answer the question. In fact, despite low numbers of questions in some question sets, all of the t tests were significant at the .01 level at least. In terms of sheer increase in responsiveness, the gains were most striking on either-or and multiple choice questions, where the percentage of the sample ultimately responsive was approximately 11 percentage points higher than the percentage responsive at first asking. The gains for picture choice and yes-no questions were moderate by comparison (3.6% and 5.4%, respectively), but this may very well have been due to the fact that responsiveness to the first askings of these questions was already almost at its ceiling. (Although roughly two-thirds of the sample appears responsive in the figure, it must be noted that these figures are for all 58 subjects, 14 of whom had failed the screening interview and thus were defined as totally unresponsive. Thus, if all interviewees who actually received the entire interview had answered a question, the responsiveness figure would have been approximately 76%.) Open-ended questions were difficult on first asking, and the gain attributable to asking the question twice (about eight percentage points) was not as large as that for the almost equally difficult verbal multiple-choice questions. By analyzing specific changes, we found that persons who had given an inadequate (vague) response the first time were most

likely to give an appropriate response the second time, but that a variety of inappropriate responses could be converted to appropriate ones if the question was repeated.

Summary and Implications

As the first challenge in interviewing mentally retarded persons is to obtain a usable answer of some kind, we devoted considerable energy to measuring and analyzing responsiveness to the communication demands of questions. The nine-category coding system we developed proved to have higher interrater reliability and to adequately classify the range of possible behaviors in response to questions. There are, to be sure, problems in coding responsiveness. For example, the judgment that a response is unintelligible is clearly a function of interviewer sensitivity; "don't know" responses under some conditions should probably be labeled appropriate rather than inappropriate responses; and there are judgments to be made in determining that an answer is too vague to qualify as an appropriate response, especially when, for example, subjects give relatively sophisticated but vacillating answers to yes-no or either-or questions which ultimately cannot be construed as taking one side or the other. Still, the responsiveness coding system generally served its purpose of providing a means of examining answering behavior independent of the reliability or validity of answers.

As indicated in this chapter, we discovered several important things about responsiveness. First, the reasons for failure to respond appropriately are diverse. In interviewing retarded persons, one can expect silence in response to questions as well as various kinds of unintelligible, vague, and totally irrelevant responses. As indicated by the predominance of minimally

appropriate responses as compared to expanded appropriate responses, one can also expect codable answers to generally be brief, in many cases monosyllabic.

As our analyses clearly demonstrate, responsiveness is a stable individual characteristic. An individual's ranking in the group based on the percentage of questions answered appropriately is remarkably stable from interview to interview. While some subjects appear to become more responsive to a second interview as a function of having participated previously in an interview, changes in responsiveness from administration to administration were slight. Moreover, responsiveness as a behavioral characteristic appears to be related to intelligence. Certainly we found that verbal interviewing techniques are infeasible with the vast majority of profoundly retarded persons. In the severe range of retardation, individual differences were very large, and it was difficult to predict which individuals would be able to respond. This suggests that screening potential subjects for their ability to participate in an interview, possibly with more stringent criteria than were used in our studies, is especially useful in the severe range of retardation. Otherwise, one can generally expect most moderately and mildly retarded persons to be able to answer most questions asked in a simple survey. Our data also indicated that while responsiveness was a function of IQ, it was unrelated to the sex, age, or place of residence (institution versus community) of the retarded person.

Perhaps the most practically important finding about the nature of responsiveness is that it depends not only on the subject's cognitive abilities but on the form in which questions are asked. Judging from the consistent findings across three

samples, we can confidently conclude that yes-no questions and questions calling for a choice between pictures are the question formats likely to optimize responsiveness. Verbal either-or questions also appear to qualify as an attractive option if obtaining codable answers from a majority of a sample is the goal. However, verbal questions requiring choice among three or four alternatives and open-ended questions appeared to be difficult to answer, particularly for the severely retarded. As we will point out in later chapters, there were other problems associated with these formats besides low responsiveness, and altogether they prove to be relatively poor sources of information in most cases. At the very least, it must be recognized that many persons in a sample will simply not be heard from when these formats are used and that their utility is greatest when interviewees have the high verbal ability required to comprehend them and generate answers. In keeping with these findings, one possible approach to survey design would be to construct layered interview schedules so that each area of questioning begins with the relatively easy-to-answer structured questions, especially those which can be answered nonverbally, and then progresses, where the interviewee's communication skills permit, to less structured questions which can then provide more specific information to clarify closed-ended responses.

Finally, our data suggest another relatively simple way of increasing responsiveness; that is, repeating questions, possibly even more than one time if necessary, in order to give retarded persons additional opportunities to understand them. The simple technique of repeating a question if the first response was less than appropriate did increase responsiveness for all types of

questions examined. If simple repetition works, what about the common practice of rephrasing questions to make them more understandable? We consciously avoided this practice in our research because we did not want to run the risk of altering responses by altering questions. As our later findings will illustrate, the way in which a question is worded or formatted can make a substantial difference in the answers that are given. Unpremeditated rephrasing of questions is particularly dangerous, for people have a natural tendency to simplify the form of the question; for example, dropping back to what appears to be a simpler and more structured format if an open-ended format does not work. Thus, when an interviewee fails to answer appropriately the first time, we would recommend starting with a verbatim repetition of the question. If that does not work, responsiveness might still be increased through the use of preplanned rephrasing of questions. However, we would recommend that these rephrasings involve changes in wording rather than changes in format simply because changes in format, while they may be more likely to increase responsiveness, are also more likely to alter the content of responses, judging from evidence to be presented later.

Finally, we might add a word about the implications of our responsiveness data for the communication development of mentally retarded persons. The predominant trend in the previous literature on the communication skills of the retarded has been to study the forms of language and the development of the ability to comprehend and produce those forms. Recently the emphasis in the study of normal language development has shifted so that researchers are more concerned with

the use of linguistic forms in various situations, with the *functions* that language serves in communication settings. The measure of responsiveness used here is a measure of behavior in a real communication setting; more precisely, a measure of an interviewee's tendency to meet the communication demands of interview questions. Although ours is not a developmental study, it suggests that there is a developmental sequence that children must progress through in understanding and answering questions. Like some previous research with normal children (e.g., Ervin-Tripp, 1970) it suggests that structured questions may be mastered earlier than unstructured questions and suggests more specifically which kinds of questions can be answered most easily by persons with low mental ages. We do not yet know the extent to which responsiveness to questions of increasing difficulty can be trained. However, we are, in the context of a related study, testing the effectiveness of such training. It would appear to us that since answering questions is a common form of verbal behavior, explicit training in responsiveness should have an important place in language arts curricula for the retarded. Such training might first aim to establish minimally appropriate responses and then encourage students to go beyond the minimal demands of the question to qualify and elaborate answers and to anticipate upcoming questions. Other research we have conducted (e.g., Elias, Sigelman, & Danker-Brown, 1980) suggests that the ability of retarded adults to answer questions appropriately is associated with making positive impressions on raters of videotaped interviews. Thus, there is some evidence that training to increase responsiveness to questions might not only contribute to communication development

but facilitate social and vocational acceptance of retarded persons.

To this point in the narrative, we have considered only the ability to respond to questions. However, obtaining an answer appropriate to the questions is only part of the battle. One might ask whether the answers

obtained are reliable and valid if one is to have confidence using interviewing techniques as a source of information about the circumstances, needs, and attitudes of retarded persons. In the next chapter, we confront the reliability issue.

TO WHAT EXTENT ARE ANSWERS RELIABLE OVER TIME?

Assuming that an answer is obtained from a mentally retarded interviewee, one must immediately be concerned with the reliability and validity of response. This chapter concerns reliability of response or the extent to which answers can be considered stable indications of interviewees' needs, circumstances, and attitudes. There are at least two ways in which the reliability issue can be approached in survey research. One approach, a variation on test-retest reliability, involves readministering the same questions after a brief interval of time and determining the extent to which answers given on the two occasions are consistent. The second approach is akin to measuring test reliability by determining the consistency of performance on alternate forms of a test. It would involve asking questions in alternative ways and determining the extent of agreement between responses to the alternative questions. In the present study, we did indeed ask a number of questions in alter-

nate forms, but we have chosen to discuss the results in the context of validity rather than reliability. Perhaps this decision could be debated, but in any event, the present chapter's discussion of reliability of response is concerned only with the extent to which answers given one week agree with answers given by the same respondents the next week.

Two interviews were administered approximately a week apart to both institutionalized children and institutionalized adults. In each sample, half of the subjects had Interviewer 1 first and Interviewer 2 second, while the other half had Interviewer 2 first and Interviewer 1 second. Since two forms of the interview schedules were used (and were also counterbalanced for order of presentation), interviewees were exposed to somewhat different interview schedules on the two occasions. Within those schedules, however, some questions remained in identical form on both occasions, and the only difference was the context of other questions.

There is no absolute criterion which can be used to make the judgment that responses are reliable or unreliable. As we shall point out, the analysis of such information is very complex, and multiple guidelines for interpreting levels of consistency may be called for. One can, in looking at the literature on test reliability, generally point to a criterion of reliability of .80 to .85 or higher, but this criterion is not typically applied to the reliability of individual test items so it serves only as a rough standard. We will indeed examine the percentage of paired responses that are consistent rather than discrepant. As we will indicate, another perspective can be obtained by examining the extent to which the degree of consistency obtained deviates from the degree of consistency which could be expected on the basis of chance, but this criterion alone is not entirely suitable either. To complicate matters further, one must consider whether the data obtained through a survey are to be analyzed at the aggregate or the individual level. It is entirely possible that many individuals might change their responses from one occasion to the next but that the overall picture of the group obtained will not change a great deal, assuming that changes by individuals essentially cancel each other out. In this same situation, however, it would be unwise to use the data to identify individual correlates of response or to make program decisions affecting individual clients.

With that in mind, let us look in some detail at the extent to which answers given by institutionalized children were consistent from one week to the next.

Reliability Over Time— Institutionalized Children

In the children's institution sample, 31 questions repeated on both forms were examined for reliability. Thirteen of these were verbal yes-no questions about activities, quality of life, and rules and decision-making. An additional four questions of the yes-no variety were accompanied by pictures illustrating the chores asked about. Table 6.1 presents a variety of information about these 17 questions and can serve to illustrate the issues in analysis of reliability.

Generally a relatively high proportion of the sample was able to respond to both askings of each of these yes-no questions, although responsiveness fell off somewhat on the rules and decision-making questions, presumably due to the difficulty of such concepts as "decide" and "allowed." Thus, the figures are based on a fairly wide range of subjects (more so than is the case for open-ended questions and other more difficult formats). For the entire set of 17 questions, an average of 86.9% of the respondents were able to respond consistently "yes" or "no" to both askings. The percentages of consistent response, as shown in the fourth column of the table, ranged from a low of 56.1% on the question about whether residents are allowed to hit people to a high of 100.00% on the question about cleaning the floor. Generally, the percentage of subjects responding to both questions who were consistent in their responses were lowest on questions to which a "yes" answer might be construed as a socially undesirable response. The questions in which a "yes" response would indicate that residents are allowed to hit people, that people yell or say mean things, and that people take things away all had low percentages of consistency.

Table 6.1: Reliability of Response of Yes-No Questions: Institution Children

Question	N Responsive	% Yes (Form A)	% Yes (Form B)	% Consistent	Kappa
Do you watch TV?	42	97.6	97.6	95.2	-.02
Do you listen to the radio or record player?	41	92.7	90.2	92.7	.53
Do you go to school or take classes?	41	95.1	92.7	92.6	.36
Do you have a family?	44	95.5	100.0	95.5	.00
Is this place your home?	40	65.0	67.5	92.5	.83**
Do people here yell at you or say mean things?	38	55.3	44.7	73.6	.48**
Do people here take things away from you?	43	44.2	48.8	76.8	.54**
Do people here help you when you want help?	39	94.9	92.3	97.4	.78**
Do people here teach you things you want to learn?	40	92.5	95.0	92.5	.36
Are you allowed to stay up late at night?	33	54.5	54.5	87.9	.25**
Are you allowed to hit people?	41	12.2	56.1	56.1	.19
Do you decide what chores you do?	24	58.3	62.5	70.8	.39*
Do you decide how to spend your money?	36	88.9	88.9	94.4	.72*
(With Pictures)					
Do you set the table?	43	97.7	93.0	90.7	-.04
Do you do dishes?	44	90.9	79.5	84.1	.38
Do you do laundry?	45	73.3	66.7	84.4	.63**
Do you clean the floor?	43	95.3	95.3	100.0	1.00*

* Significant at the .05 level (one tailed) ** significant at the .01 level.

Consistency for the question about decision making regarding chores was also low, possibly in this case because two difficult concepts, "decide" and "chore," may have created confusion. Generally, consistency figures for the 13 verbal questions (average consistency 86.0%) were similar to those for the questions accompanied by pictures (average consistency, 89.8%). Generally the predominant form of consistency on most of these questions (with the exception of the negative items discussed above) was saying "yes" both times rather than saying "no" both times. "Yes" answers predominated on most of these questions. Moreover, the percentage of subjects saying "yes" on Form A ($r = .88$) as well as on Form B ($r = .77$). This suggests that reliability is artificially inflated when marginal distribution is lopsided. As will be noted later, acquiescence was common in our samples; thus we suspect that the reliability figures for yes-no questions are generally inflated. We would at least warn that the relatively high consistency figures in Table 6-1 are not necessarily evidence of validity. However, we cannot estimate precisely the extent to which acquiescence might have contributed to consistency.

In any case, the figures are encouraging in the sense that we can expect much the same response both times from most subjects most of the time. As the second and third columns of the table indicate, this means that the pictures of the group obtained from two askings of the same questions are highly similar, the biggest discrepancies appearing on the question about being allowed to hit people and on the picture questions about dishes and laundry, both of which were associated with only moderately high consistency figures.

A different perspective on reliability emerges from the last column

of the table, which reports the kappa statistics calculated from each 2 x 2 contingency table analyzed. Kappa is a straightforward estimation of the extent to which the obtained percentage of consistency between paired responses exceeds the percentage which might be expected based on chance, chance probability being calculated from the marginals of the contingency table. Assume, for example, that in the sample 90% actually watch TV and 10% do not, and that 90 and 10 are the marginal percentages both times we ask the question. Now assume that we randomly pair first responses and second responses rather than pairing two answers given by the same respondent. As discussed in Chapter 3, we can calculate the probability of agreement between pairs of response by chance from the marginals, multiplying $.9 \times .9$ and $.1 \times .1$ and adding the two to yield a probability of .82. This means that 82% of the pairs of response would agree simply by chance. When response distributions are lopsided like this (e.g., when there is not approximately a 50-50 split between "yes" and "no" responses at each asking), the chances of agreement become relatively high as illustrated in this example. Thus the obtained percentage of agreement, even if it is relatively high, may be no higher than what we would expect if we simply paired by random assignment one subject's response with another subject's response instead of pairing two responses given by the same subject. (For those whose memories of statistics are jogged by the image of red and green balls in a jar, the probability of drawing two red balls becomes progressively higher as the number of red balls as compared to green balls in the jar increases.) In interpreting kappa statistics then, we must bear in mind this question:

To what extent are people responding so consistently to the question both times that their consistency is greater than we would expect by chance? As can be noted in the table, several questions which have very high consistency figures do not have large or statistically significant kappas, while several questions with much more modest consistency figures do have significant kappas. When a question yields an extremely high proportion of "yes" responses on both askings, it becomes very difficult to obtain a large kappa simply because the odds of agreement by chance alone are also high.

We are not willing to say that kappas should be used as the ultimate criterion of reliability. We consider it impressive to obtain 90% or greater consistency figures even if those figures are not a significant departure from chance expectations. Such high figures indicate that the picture of the group is reliable and that individuals within the group are saying essentially the same things on two occasions a week apart. On the other hand, we can be even more satisfied when the consistency figure and the kappa associated with it are high, as in the last question in the table about cleaning the floor. Similarly, we can be particularly suspicious of the utility of a question if the consistency figure is low and the kappa is low as well, as is the case for the notably ineffective question about whether residents are allowed to hit people. In between, we have softer ground to stand on and must take into account all the information obtained, the percentage saying "yes" on each occasion, the percentage of responses that are consistent, and the extent to which this percentage deviates from chance expectation.

Overall, we would conclude that yes-no questions appear to yield fairly

reliable data. However, we would reemphasize the warning about acquiescence, because saying "yes" on two occasions does not necessarily mean that "yes" is the valid response.

Six multiple choice questions repeated across the forms also posed interpretive problems. Three of these questions which asked subjects to identify the extent of their activity as "a lot," "sometimes," "not much," or "never" (or in one case "a lot," "some," "not many," and "none" in reference to number of friends) yielded 71.9% consistency on the average. This figure is relatively high considering the fact that the probability of matching responses by chance is lower for a four-choice question than for a two-choice question. Indeed the kappas for these items ranged from .51 to .64 and were all significant at the .01 level. However, three other questions, asking respondents how they felt about aspects of institutional life, yielded extremely low consistency (46.2 on the average), even though the kappas for two of the questions (.27 and .20) were statistically significant. It is impossible to determine whether this was the case because of the subjective and mood-dependent nature of the questions or because a greater proportion of the sample was able to respond to these pictorial questions than to the verbal multiple choice questions. In any case, this format cannot be considered successful.

The remaining eight questions repeated on both forms were open-ended. Four were factual in nature, calling for name, spelling of name, birthday, and characterization of the state school. Roughly two-thirds of those asked these questions were responsive to the first and second administrations of each; of these 63.3% were consistently able to give their first and last names correctly; 42.9% were consistently able to spell

their names; 63.6% consistently identified their birthdays; and 61.9% consistently described the place in which they lived. This finding is rather discouraging, not only because many persons could not answer these basic questions, but also because those who did often gave incorrect information both times. On the other hand, the criteria for correctness were rather strict.

The remaining four open-ended questions were each analyzed by response categories, which were based on responses given by both residents and attendants. In these cases, both the total percentage consistent (i.e., either mentioning an item in a category on both askings or failing to mention it both times) and the percentage mentioning something both times were examined. There were marked differences between the two measures. On a question asking for an enumeration of activities with friends, for example, the average total consistency figure across all eight categories was 70.7%, but the average percentage of respondents consistently mentioning something both times was only 18.1%. Virtually the same thing occurred for the other questions asked. For example, asked about what they were being taught in school, 69.6% of those who could answer both times referred to academics of some kind, but for each other category of response, high total consistency was achieved by failing to refer to a category both times. There seems to be no satisfactory way of analyzing such questions, for consistency depends on both the verbosity of the respondents and the number of categories developed by the coders. Perhaps the major conclusion to draw is that open-ended questions calling for enumerations yield very little information, partially because many mentally retarded persons cannot respond appropriately to them in the first place, and par-

tially because when they do respond, they say very little. Due to the predominance of "no mentions," only 9 of 35 kappas were significant. However, since the total consistency figures for all 35 response categories of four questions averaged 83.9%, one can at least conclude that it was relatively rare that a respondent mentioned something one time but failed to mention it the other time.

Judging from this sample, reliability of response from one week to the next can generally be considered marginally adequate, except for the subjective yes-no questions mentioned previously about rules and decision making and the picture-choice questions about the quality of institutional life. Reliability did vary as a function of both question format and question content, but for the most part answers to questions repeated a week apart were largely similar.

Reliability Over Time— Institutionalized Adults

In the adult institution sample, 48 questions were repeated on both forms and could be examined for reliability. The questions were analyzed in three separate groupings according to question format: yes-no, either-or, and open-ended.

The 29 yes-no questions were divided into those which elicited factual information and those which required a subjective response from the interviewee, with the expectation that reliability might be higher for factual questions. Analysis of response consistency across both survey forms revealed that the average consistency for yes-no questions was 81.2%. Contrary to expectation, there was almost no difference in consistency between the factual and the subjective questions (80.6% and 81.7%, respectively). Response con-

sistency across this set of questions ranged from 97.3% to 60.5%. The highest consistency figures were found for questions like "Are you usually happy?" "Do you have a family?" and "Do you know how to cook?" while the lowest consistency figure (60.5%) was obtained on the question "Are you usually sad?" another question for which "yes" can be considered a socially undesirable response.

There was a definite tendency for the majority of interviewees to consistently give "yes" responses to many questions (21 of 29). It appeared that the reliability figures for this sample, as was the case for institutionalized children, were inflated by a tendency to acquiesce on yes-no questions. Reliability was highest for those items for which "yes" was a socially desirable choice. However, even if such a response set were not a significant factor in the high reliability figures obtained, it cannot be assumed that consistent responses to these yes-no questions necessarily represent accurate or valid information.

Thus the picture of this sample based on their responses to these questions appears to be primarily positive, with the majority of the interviewees seeing themselves as usually happy or possessing a number of basic adaptive skills such as reading, writing, keeping house, counting money, and cooking. This picture may have been influenced by a tendency to acquiesce or to give socially desirable responses as already discussed, and indeed we will demonstrate later that attendants are not so ready to agree that clients have basic adaptive skills. However, the responses to questions such as, "Are you usually sad?" and "Do you have any problems?" seem to suggest that many clients can say "no" to deny having difficulties.

There were four verbal either-or questions repeated on the two alternate surveys consisting of two pairs of questions with the order of options varied: "Are you usually happy or sad?" vs. "Are you usually sad or happy?" and "Are you usually by yourself or with other people?" vs. "Are you usually with other people or by yourself?" Average consistency for these four questions was again relatively high (82.2%) and was comparable to the average consistency obtained for yes-no questions. A substantial majority of respondents consistently indicated that they were usually happy in response to both happiness questions. However, there was a tendency for more interviewees to consistently answer "sad" when this alternative was given second (12.0%), than when it was the first option (2.9% consistently reporting "sad").

The relationship between the two questions of the second either-or pair was more complicated, although a similar pattern emerged. The majority of respondents consistently replied to both questions that they were usually with other people as opposed to being by themselves. As with the previous question pair, the percentage who consistently chose the alternative "with other people" was slightly higher when this choice was presented as the second option in the question than when it was first (50.0% vs. 46.7%). In addition, the percentage who consistently chose the "by yourself" alternative doubled when this option was presented last (33.3% vs. 16.7%). For this question, "Are you usually with other people or by yourself?", the effect of consistently choosing the last alternative was to increase the overall percent of respondents who answered the question consistently both times to 80.0% for this question as compared to 66.7% for the alternative question. The observations made from these two either-

or question pairs seem to indicate that a small portion of interviewees are probably influenced more by the order of the either-or alternatives than by the actual content of the response alternatives.

The last 15 questions on both surveys were open-ended. The first seven open-ended questions asked the residents to give factual information concerning their names, birthdates, and current addresses. Responses to these questions were coded either correct or incorrect based on records kept on the subjects. On the average, 51.9% of the residents who responded to the questions consistently gave correct answers. Residents were most likely to give correct information about their dates of birth: 77.4% correctly identified the month they were born in, 72.2% consistently gave the correct date, and 58.8% were able to give the right year both times. In comparison, only one-third (33.3%) of the residents could correctly give their full present address consistently, while slightly more than half (56.7%) were able to correctly give their first and last names at both interviews (only first or last names were scored as partially correct). Only 25.0% of the residents could consistently spell their names (orally) when asked, making this the most difficult question of the seven. A greater number of subjects (40.0%) were able to correctly write their names both times. Five of the 15 were coded with present response categories, and the average consistency for these questions was 83.3%. This figure would have been higher except that response consistency for one question, "What kind of place is this?", was exceptionally low, with 55.0% giving consistent descriptions and only 30% consistently identifying the place where they lived as a state school. A clear majority of the respondents consistently indicated that someone

other than themselves made the rules where they lived (92.0%) and decided what chores they would do (75.0%). In contrast, the respondents were almost equally divided on the question "who decides how you spend your money?", with 50.0% consistently answering "someone else" and 43.8% consistently indicating that they made the decision themselves. Total reliability figures for these three questions were uniformly high: 96.0%, 83.3%, and 93.8%, respectively. These findings indicate that most subjects who can answer give reliable answers to open-ended questions concerning who makes decisions about their everyday activities. In addition, 90.8% of the subjects consistently answered the question, "How many people sleep in your bedroom?"

The remaining three open-ended questions assigned each resident's answer to one of several response categories made up on the basis of the responses given by both residents and attendants. The average consistency for this set of 26 open-ended response categories was 87.4%, the highest average consistency for any of the question groups analyzed. However, the consistency percentages for the response category questions were inflated by the large number of categories whose high consistency was achieved by clients' repeatedly failing to mention anything in the category. For example, on one question which asked the resident to name the person who helps him/her the most, there was only one category (cottage staff) for which the most common response combination was mentioning that category both times. For the other eight response categories for this question the most common response combination was not mentioning the category either time. Therefore, while the average total consistency figure across all nine response categories was high (81.5%), the average percentage of respondents

consistently mentioning something both times was very low.

This same pattern was found for the remaining two response category questions. When residents were asked who they usually talked to about their problems, the most consistently mentioned response category was "cottage staff" (36.4%). Four out of the remaining six categories were mentioned by only a few residents, making the average percent of respondents mentioning a category both times again quite low (9.1%). On the final question of this group, "If you had one wish, what would you wish for?" no response category was mentioned very frequently on both occasions. The most frequently and consistently mentioned wishes were to be with their families (27.3%) and to have material items (13.6%). Consistency in responding to this question was achieved primarily by failing to mention a response category both times. Thus open-ended questions, which could not be answered by many subjects in the first place, were reliably answered, but rarely generated consistently mentioned information.

Overall, for this sample of institutionalized adults, response consistency for questions repeated at a one-week interval averaged over 80%, providing interviewers with answers that were reasonably reliable. There was little variation in response consistency associated with the three different question formats. However, there was some indication that high consistency on yes-no questions may have been gained in part through acquiescence, that some subjects prefer the second alternative on either-or questions, and that consistency on open-ended questions was gained primarily by failing to mention something both times rather than mentioning something twice.

Individual Correlates of Response Reliability

As was done in analyzing responsiveness as an individual characteristic, an analysis was conducted to determine what accounts for individual differences in reliability of response from one week to the next. In the children's institution sample, a reliability score was formed based on the number of instances of consistent responses to 17 repeated yes-no items divided by the number of item pairs answered appropriately. This index was calculated only if subjects had been able to answer half or more of the 17 repeated item pairs. A comparable index was formed for institutionalized adults based on 29 yes-no items. Then, in each sample, this reliability score, which could theoretically range from 0% to 100%, was correlated with IQ, age, sex, responsiveness to the entire Form A interview schedule, and an acquiescence index, on which subjects were given a score of 1 as opposed to 0 if they said "yes" to both "Are you usually happy?" and "Are you usually sad?" We expected that subjects with higher IQs and higher responsiveness scores indicative of verbal skill would be more likely than lower IQ and less responsive subjects to give consistent responses on two occasions. At the same time, we predicted that subjects who tended to acquiesce might be likely, on yes-no items at least, to achieve high reliability scores simply by saying "yes" a great deal. No specific predictions were formulated about the relationship between age and sex and reliability of response.

Reliability scores in both samples were normally distributed. In the children's sample, these scores ranged from 59% to 100% (mean = 85.8%, SD = 11.1%) for the 44 of 52 subjects for whom they were calculated. Similarly, in the adult sample, 40 of 58

subjects had enough data to calculate scores, and scores ranged from 58% to 97% (mean = 76.1%, SD = 10.2%). Interestingly, then, no one was inconsistent on more than half of the items for which first and second administration answers were available.

Table 6.2 presents the simple correlations between individual characteristics and the reliability of response score. These correlations are very difficult to interpret, particularly since the findings in the two samples do not correspond well. In the children's institution sample, contrary to expectation, lower IQ subjects were actually more likely to give reliable responses than higher IQ subjects. Other correlations were small and inconsistent in direction, but there was at least a tendency for more acquiescent subjects to display more reliability of response. In the adult institution sample, by contrast, there were small positive associations between both IQ and responsiveness and reliability, the relationship between responsive-

ness and reliability being statistically significant. In this sample, there was essentially no relationship between acquiescence and reliability. About the only consistency across samples was a very slight tendency for females to provide a greater proportion of reliable responses than males did.

These simple correlations do not take into account the fact that some of the individual characteristics involved in the analysis are inter-related. As noted in Chapter 5, IQ and responsiveness were positively associated in both samples. Moreover, there is a consistent tendency for higher IQ persons to be less likely to acquiesce as indicated by responding "yes" to both happiness items. The correlation between IQ and acquiescence was .40 in the children's sample, and .31 in the adult sample. Therefore, partial correlations, which allow for examining a relationship between two variables with the effect of a third controlled, were conducted to attempt

Table 6.2: Individual Correlates of Reliability of Response

	Child Institution Reliability	Adult Institution Reliability
IQ	-.38**	.12
Age	.06	.22
Sex	.12	.19
Responsiveness	-.15	.32*
Acquiescence	.16	-.10

*Significant at the .05 level

**Significant at the .01 level

to unravel the independent impacts of IQ, acquiescence, and responsiveness on the reliability measure. However, these analyses did not clarify the findings. They left the pattern of correlations in Table 6.2 essentially unchanged.

Thus, we must conclude that the extent to which mentally retarded persons give consistent answers on two occasions is essentially unpredictable based on the information we had available. It is still possible that opposing forces were at work; that, on the one hand, higher ability subjects were likely to be consistent in their responses but that, on the other hand, acquiescent tendencies inflated consistency scores for lower ability subjects. Unfortunately there were not enough items in formats other than yes-no to allow us to see if reliability of response where acquiescence does not operate is more predictable.

Summary and Conclusions

This chapter has been concerned with one aspect of the issue of whether answers given by mentally retarded persons can be considered useful information. In both institution samples, repetitions of questions a week apart provided evidence of response reliability. In both samples, it was found that answers given on two occasions were generally consistent over 80% of the time. Figures for yes-no, either-or, and open-ended questions with response categories were generally very similar and high. The lowest consistency figures were obtained for the multiple choice questions used in the children's institution sample, especially those questions about satisfaction which asked respondents to choose among happy and sad faces. At the same time, as kappa statistics associated with these items suggested, it is more difficult to obtain con-

sistency by chance on four-option questions. Thus, although the consistency figures for these items were relatively low, it appeared that respondents were not just answering in a random way.

While we might be more gratified by consistency closer to 100%, we must conclude that what mentally retarded persons say one week is likely to be quite similar to what they say the next week. Most of our questions were factual in nature so such consistency should be expected. Interestingly, however, an analysis in the adult institution sample of factual versus subjective yes-no questions did not reveal lower reliability of responses to subjective questions, despite the fact that such responses might be predicted to be more mood-dependent and changeable from week to week.

As we have noted throughout this chapter, however, reliability of response figures was very difficult to interpret. In response to yes-no questions, an unthinking endorsement of items would be scored as consistency of responses in exactly the same way as a thoughtful provision of "yes" responses on both occasions. Given our findings of high rates of acquiescence to yes-no questions in retarded samples, as well as the positive association between the percentage of the sample endorsing an item and the consistency figure for that item in the present analysis, we must conclude that reliability figures for yes-no questions are inflated. At the very least, we must be cautious not to construe them as evidence that answers given a week apart express a consistent body of "information" about the lives of the respondents. We can conclude that the picture of the group, and of the individual within it, obtained from yes-no questions does not change much from week to week, as long as we recognize that these

"pictures" may not be faithful representations of reality.

Much the same can be said of reliability figures for open-ended questions. We suspect that answers to them underrepresent the extent to which mentally retarded people are, for example, involved in various activities, in the same way that yes-no questions probably overrepresent extents of involvement. While we did not have a completely satisfactory way of analyzing consistency for these questions, it was apparent that few responses were given to such questions and consistent mention of items was a rare phenomenon simply because mention of items was a rare phenomenon. Unfortunately, the reliability of such items was heavily influenced by the number of categories that the coders chose to construct. The fact that reliability for categories with frequent mentions tended to be lower than reliability for categories with virtually no mentions is not a healthy sign.

As for items with structured choices, they were associated with relatively low reliability figures. While either-or items fared well despite the fact that they were subjective in nature, verbal multiple choice items calling for an indication of extent of activity were only marginally reliable. Pictorial multiple choice questions about satisfaction with aspects of institutional life were

not responded to in the same way on two occasions, even though obtained consistency figures exceeded chance probabilities. These figures, too, are difficult to interpret. Subjective feelings about the institution might well change from week to week as a function of mood and intervening experience, even "true" answers are given given on both occasions.

All in all, then, our analysis of reliability is encouraging in the sense that answers can generally be counted on to be the same from week to week, but discouraging in the sense that we are not sure how to interpret even high reliability figures. Moreover, we were unsuccessful in identifying which persons are most likely to provide reliable data. This uncertainty will, however, be reduced as we turn to issues of response validity. In the next chapter, we will take up the question of whether clients' responses agree with responses to the same questions provided by attendants and parents. In this analysis, we will be able to look more directly for evidence of systematic biases in the responses given by mentally retarded persons. Such biases as acquiescence, selection of the last option on either-or or multiple choice questions, and failure to mention things mentioned by significant others, to the extent that they occur, may cast the reliability figures presented here in a different light.

TO WHAT EXTENT ARE ANSWERS VALID?

The two previous chapters have dealt with responsiveness of mentally retarded persons to interview questions and with the reliability of their responses. While it is clear that these factors are important, it is equally clear that even reliable responses to questions are useless if they are not valid. That is, it is pointless to elicit responses which do not accurately reflect the attitudes and behaviors of the mentally retarded persons interviewed. We have devoted considerable attention to evaluating the validity of responses by our mentally retarded respondents.

Our main strategy for assessing validity has been to compare the responses of clients with the responses of adults who should be knowledgeable about the mentally retarded interviewees. Thus, when we asked a child "Do you go to church?" we asked the child's parent or cottage attendant if the child goes to church and noted the per-

centage of cases in which the two answers agreed. Clearly, there are instances in which the parent or attendant will not be in a good position to provide information regarding the mentally retarded respondent. Especially when the attitudes of retarded persons are at issue, attendants' or parents' opinions may not be accurate or even meaningful. Then, too, we must recognize that even on points of fact parents, and perhaps especially attendants, may simply not be knowledgeable, and their answers, like those of retarded persons, may be biased. Nonetheless, we assume that the degree of correspondence of the responses of children and their significant others is a useful (although admittedly imperfect) indicator of response validity.

A second strategy for evaluation of response validity has been to ask questions to which the correct answer is known. This technique reduces the ambiguity inherent in the first

strategy. If, in response to such questions, mentally retarded persons are able to provide accurate information, we have sound evidence of their ability to give accurate responses to at least some interview questions.

Our use of these techniques in the course of this project was extensive. Therefore, a discussion in detail of all validity data would be impossible. Fortunately, the analyses of agreement in the different samples yielded strikingly consistent results. Thus, we will discuss in detail the comparisons of children's samples, using that data to illustrate the major conclusions. Subsequently, agreement data from the additional samples will be discussed, citing only those data which either contradict the conclusions derived from the community children's sample or which add novel insights. Then, the validity data derived from the use of questions to which answers are known will be discussed. Finally, we will look for differences among mentally retarded persons in their tendencies to give valid responses as indicated by agreement with significant others.

Community Children

OVERALL AGREEMENT

Sixty-one questions were asked of both community children and their parents, and the degree of correspondence between their responses was evaluated as a check of response validity. Average agreement for this set of questions was 64.5%; that is, parents and their retarded children were in agreement about two-thirds of the time. Overall agreement and a breakdown of agreement according to question type are depicted in Table 7.1. Clearly agreement did vary as a function of question type, but as the discrepant figures for two different types of

multiple choice questions suggest, it is necessary to look more closely at both question format and question content in order to understand what makes agreement likely to be high or low.

YES-NO QUESTIONS

Yes-No Questions With Words Only

Average Agreement

Questions posed of non-institutionalized children and their parents included 22 simple verbal questions, most about client activities, requiring "yes" or "no" responses. The responses of clients and their parents were compared; data for these 22 comparisons are summarized in a mean contingency table (Table 7.2).

Children and parents tended to agree in their responses to yes-no questions; on the average, 65.7% of comparisons yielded child-parent agreement. As can be seen, agreement tended to result primarily because both clients and parents responded "yes." Because clients favored "yes" slightly more than did their parents, the majority of disagreements occurred when the client said "yes," and the parents said "no," suggesting that acquiescence by retarded persons may threaten the validity of their answers.

Differences in Agreement Due to Question Topic

The use of averaged data can obscure important factors underlying responding. In fact, agreement varied greatly from question to question, ranging from 92.8% ("Do you watch TV?") to 40.8% ("When you're not at school, are you usually by yourself?"). In pursuit of explanations for this variation, we examined in some detail the differences in agreement rates within this group of questions. First, it would appear that relatively

Table 7.1: Agreement by Question Type 2

Question type	# of Questions	% Agreement
Yes-no (words only)	22	65.7
Yes-no (with lead-in & pictures)	10	71.5
Either-or (subjective; words only)	7	65.6
Either-or (subjective; with pictures)	4	58.1
Multiple choice (discrete)	6	77.9
Multiple choice (quantitative)		24.0
Open-ended (factual)	4	64.8
* Open-ended (coded responses)	3	77.9
Total	61	64.5

*Calculated by averaging the average agreement scores across response categories for the three questions

simple and concrete questions yield higher agreement rates than do more complex and abstract questions. For example, agreement for "Do you

go to church?" was 78.2%, as opposed to 56.4% for "Most of the time, is it up to you to decide what time you go to bed?"

Table 7.2: Mean Contingency Table for 22 Verbal Yes-No Questions

Responses of Clients	Responses of Significant Others		
	No	Yes	
No	13.0%	14.1%	27.1%
Yes	20.1%	52.7%	72.8%
Total	33.1%	66.8%	

Agreement = 65.7%

Five questions which asked about the client's participation in various activities had an average agreement rate of 82.3%, in contrast to 66.5% for a group of seven more complex and abstract questions dealing with training in specific skills ("Is anybody at school teaching you about money now?" is an example of the latter group of questions).

Second, it appears that some topics include built-in response biases which influence agreement. For example, three questions which inquired about client's possession of particular skills ("Do you know how to read books?") resulted in agreement lower than that for any other topic area--54.3%. It seems unlikely that this is attributable to the difficulty of these questions; rather, a specific bias appears to be in operation here. That is, clients frequently claim skills which their parents deny they possess. On the average, 81.1% of clients and only 52.7% of parents answered "yes" to the skills questions, so that fully 37.0% of response comparisons were disagreements in which clients claimed skills and were contradicted by their parents.

Is this a matter of client's overestimating their skills or parents underestimating the skill levels of their mentally retarded children? Although we cannot be certain, the former possibility seems more likely in this case. In responding to skills questions, clients probably become concerned with favorable self-presentation so that a desire to appear competent encourages an affirmative answer. "Yes" answers probably also reflect a more general tendency toward acquiescence, but concern with self-presentation is probably what accounts for the particularly low agreement on these skills questions.

Differences in agreement as a function of parents' responses

Another pattern appeared in variations of agreement rate within the 22 yes-no questions. Specifically, agreement was highest for those questions to which the parents are most inclined to respond "yes." This is to be anticipated in light of an observation made throughout the course of this project that mentally retarded respondents tend to acquiesce, or to respond "yes" regardless of question content. To the extent that acquiescence on the part of the clients occurs in response to yes-no questions, higher agreement should be observed on those questions for which the response of significant others is "yes." Here agreement would not indicate response validity as much as a happy coincidence of response bias and reality.

To test that possibility, this group of questions was divided into the 11 questions which parents tended most often to answer "yes" (average of 88.9% "yes" responses by parents) and the 11 questions to which parents were least inclined to respond "yes" (average 44.6%). Comparison of the agreement rates for these two groups of questions is depicted in Table 7.3. As can be seen, average agreement for "high yes" questions was 75.7%, as opposed to 55.6% for the "low yes" questions. Looked at another way, the percentage of parents responding "yes" to these 22 questions correlated .80 with agreement figures for the questions. This comparison suggests that clients tend to indiscriminately respond "yes" to yes-no questions, and that client-parent agreement is to some extent artificially controlled by the proportion of "yes" responses given by parents. Thus overall agreement rate provides an inflated estimate of validity when the questions involved are predominantly answered "yes" by significant others.

Table 7.3: Child-Parent Agreement as a Function of Parent's Tendency to Say "Yes"

11 Questions, Low Proportion of Parent "Yes" Responses			
Client	Significant Other		
	No	Yes	Total
No	22.6%	11.6%	34.2%
Yes	32.7%	33.0%	65.7%
	55.3%	44.6%	
Agreement = 55.6%			
11 Questions, High Proportion of Parent "Yes" Responses			
Client	Significant Other		
	No	Yes	
No	3.3%	16.5%	19.8%
Yes	7.6%	72.4%	80.0%
	10.9%	88.9%	
Agreement = 75.7%			

A set of questions to which 50% of parents' responses were "yes" would provide a better test of validity of yes-no questions.

Conclusions

In short, these data suggest three conclusions. First, cognitively demanding questions posed of

mentally retarded individuals tend to elicit less valid responses than do easier questions (to the extent that comparison with their parents' responses constitutes an adequate measure of validity). Second, specific topic areas (for example, questions about skills and competencies) may involve biases which tend to

systematically alter responses in such a way as to seriously reduce validity. Third, average agreement rate is an overestimate of validity for this group of questions, for which the parents' response was "yes" 66.8% of the time. Where parents respond "yes" with frequency, acquiescence on the part of retarded respondents inflates agreement figures but not necessarily validity. It is likely, in light of this observation, that the actual validity of these yes-no questions is low enough that their usefulness as a means of gathering information from mentally retarded persons is questionable.

Yes-No Questions with Lead-in and Pictures

Interviews with community children included 10 yes-no questions which were preceded by a verbal lead-in and which used pictures to clarify questions. These 10 questions concerned chores, and the verbal lead-in read: "Here are some pictures of different kinds of chores some people do. I want you to tell me if you do any of these chores at

home." The questions were simple yes-no questions such as "Do you set the table?" and "Do you make beds?" each question being accompanied by a picture of youngsters engaging in the chore involved in the questions.

Average agreement

Data for the 10 questions were averaged and are depicted in a mean contingency table (Table 7.4). Average agreement for these questions was 71.5%. Nearly two-thirds of the response comparisons were agreements in which both child and parent responded "yes." Inconsistencies were divided between cases in which the client responded "yes" and the parent "no," and the converse; thus, there was not strong evidence of acquiescence for the set as a whole. In short, agreement for this class of questions was relatively high, somewhat higher than for the verbal yes-no questions, and tended to result because both client and parent responded "yes" to most questions.

Table 7.4: Mean Contingency Table for 10 Yes-No Questions with Lead-in and Pictures

Responses of Clients	Responses of Significant Others		
	No	Yes	TOTAL
No	6.9%	14.7%	21.6%
Yes	13.7%	64.6%	78.3%
	20.6%	79.3%	

Agreement = 71.5%

Differences in agreement

There was relatively little variability in agreement among these questions. For 9 of the 10, the range of agreement was only from 65.5% to 85.5%. It may be argued that the uniformly high agreement for these questions results from the fact that the "valid" answer to most of them (as reflected by the parents' response) was "yes." However, agreement for "Do you cook on the stove?" was only 47.2%, by far the lowest agreement for the set. It is not coincidental that this question also had the lowest percentage of "yes" responses from parents (52.7%). Clients tended to respond "yes" to this question (65.4%), with the result that 32.7% of the comparisons were inconsistencies in which clients responded "yes" while their parents were responding "no." This interpretation is further supported by a comparison with the question from this group which had the highest proportion of "yes" responses from parents: "Do you pick up stuff around the house?" Ninety-six percent of parents answered "yes," and the agreement rate was 85.5%, highest for this set of 10 questions. All agreements in this case resulted because both client and parent responded "yes." For this group of questions, the correlation between proportion of parents responding "yes" to a question and agreement for that question was .97.

Conclusions

The high average agreement for these yes-no questions appears to overestimate response validity. Average agreement is high because the apparently appropriate answer to most of these "chores" questions is "yes." When parents give a high proportion of "no" responses, client-parent agreement drops accordingly, as it did on verbal yes-no questions.

EITHER-OR QUESTIONS

Either-or Questions with Words Only

Seven questions asked of community children and their parents were simple verbal questions posing two response alternatives. "Are you usually happy?" for instance, is the complement of the above example.

Average Agreement

Agreement for this group of questions ranged from 50.0% to 86.0%; average agreement was 65.6%. Average agreement was thus comparable to that for yes-no questions using words only (65.7%), but the range for those yes-no questions was much greater (40.8% to 98.2%). To what may this decreased variability for the either-or questions be attributed? As has been pointed out, the upper end of the distribution of agreement rates for yes-no questions appears to be an overstatement of the validity of those questions, due to acquiescence by mentally retarded respondents and the fact that the parent's response to those questions is typically "yes." It seems likely that agreement rates for either-or questions are more uniform because no such biases operate in these questions to generate some spuriously high estimates of validity. Indeed, we have found no reason to suspect that agreement rates for either-or questions are anything other than an accurate indication of the validity of responses to these questions, for disagreements were generally mixed in nature. Despite roughly equivalent agreement rates for the two types of questions, then, it seems likely that either-or questions generate responses of greater validity. Agreement for yes-no questions is inflated simply because our survey unfortunately included a predominance of questions which were commonly answered "yes" rather than "no."

Differences in agreement

No difference in agreement due to the reversal of order of the response alternatives was apparent. The rates of agreement for two of the three pairs did not differ (50.0% vs. 51.0% in one pair, and 72.5% vs. 73.2% in the other). In the third pair, the "happy or sad" example stated above, agreement figures for the two forms were 86.0% (with "happy" first) and 72.0% (with "sad" first). No explanation for this difference is readily apparent. Where we have observed order effects on either-or questions in our interviews, they have typically taken the form of endorsement of the last-mentioned alternative by mentally retarded respondents. That bias, however, does not account for the agreement difference observed in this pair of questions. In fact, more clients said they were usually happy when "sad" was the last mentioned alternative. It would appear that agreement for these either-or questions does not differ systematically as a function of the order in which alternatives are presented.

There do, however, appear to be differences in agreement due to question topic, just as there were for yes-no questions. While it is difficult to predict which topics will yield agreement with significant others, variability in agreement as a function of question topic at least suggests the need for caution in interpreting data gathered in interviews with mentally retarded persons, and also suggests that we be conservative in generalizing demonstrations of validity across content areas.

Either-or Questions With Pictures

The group of questions asked of community children and their parents included two pairs of either-or questions which used pictures as the response alternatives. The members of each pair involved the same con-

tent but reversed the order in which the response alternatives were presented. For example: "This picture (point) shows a boy/girl who is happy, and this one (point) shows a boy/girl who is sad. Which picture shows how you usually feel? Point to the picture."

Average agreement for these either-or questions with pictures was 58.1%. Agreement for the individual questions ranged from 41.5% to 74.1%. Again, it appears that agreement varies more or less unpredictably across topic areas. Agreement was much higher for the two happy-sad questions (74.1% and 66.7%) than for the other two questions, which asked whether the client spends free time alone or with others (41.5% and 50.0%). Disagreement on the latter pair of questions tended to result because clients chose the "alone" pictures, while parents responded "with others." The agreement rate for these four questions did not vary systematically as a function of the order in which response alternatives were presented. Moreover, the use of pictures as response alternatives did not improve agreement over that for verbal either-or questions on the same topics. If any difference can be said to exist, it is that either-or questions with words yield somewhat superior agreement (64.8% vs. 58.1%). This apparent decrease in validity for picture questions may be due to the fact that the use of pictures increases responsiveness. Many individuals who do not respond verbally can respond by pointing to pictures. The response patterns of these less verbal, lower IQ respondents may be largely random, reducing validity. We do in fact have some evidence that the responses of lower IQ respondents tend to be less valid (see discussion at end of chapter).

Conclusions

Average agreement for either-or questions is similar to agreement for yes-no questions. Because agreement figures probably give an overestimate of the validity of yes-no questions, we can conclude that either-or questions generally yield more valid responses. There is no evidence suggesting that agreement is not an accurate indicator of validity for either-or questions. Agreement does not appear to be affected by reversing the order of presentation of response alternatives, although it does vary as a function of question topic. Finally, the use of pictures as response alternatives apparently decreases validity of responses to either-or questions somewhat, probably because less capable clients can then provide answers.

MULTIPLE CHOICE QUESTIONS

Eleven questions posed of non-institutionalized children and their parents were multiple choice questions, involving a choice between either three or four response alternatives. Average agreement was 57.0%. These questions differed in question format: some were three-choice and some four-choice questions, some used words only and some words plus pictures in depicting response alternatives, and some of the questions were follow-ups to yes-no questions. However, there were no systematic differences in agreement due to any of these format considerations.

Agreement for multiple choice questions did, however, differ according to the type of response alternatives. Six multiple choice questions gave discrete alternatives about matters of fact ("Do you live in a house, an apartment building, a trailer house, or a duplex?"), while five used quantitative dimensions as response alternatives ("How many friends do you have: a lot, some, not many, or

none?"). Average agreement for the discrete questions was 77.9% (range 70.5% to 85.5%); average agreement for the quantitative multiple choice questions was only 24.0% (range 25.9% to 39.0%). As is readily apparent, not only does average agreement for the two groups of questions differ, but their ranges do not come close to overlapping. Clearly, agreement was superior for the questions offering discrete alternatives. Because quantitative responses require finer discriminations, this trend might well be present in normal populations as well. However, the fact remains that average agreement was very high for the discrete questions and unacceptably low for those offering quantitative alternatives. This suggests that discrete multiple choice questions may be useful in interviewing mentally retarded persons, but it would appear pointless to ask mentally retarded respondents to make quantitative judgments, given the absence of evidence for validity of their responses.

Conclusions

Multiple choice questions offering discrete response alternatives yield very high client-parent agreement. Agreement for questions using quantitative alternatives is so low as to make these questions useless in obtaining information from mentally retarded respondents. No conclusion can be offered regarding format differences such as number of options and method of presentation of response alternatives.

OPEN-ENDED QUESTIONS

Open-ended Factual Questions

Questions allowing comparisons of the responses of children and their parents included four open-ended questions having a single correct answer. On the average, children matched their parents 64.8%

of the time. It is somewhat discouraging that validity was this low for simple questions requesting factual information which should be readily available to mentally retarded respondents. Agreement was only 42.5% for the question "Counting you, how many people live in your house right now?" The other three questions asked the client's age, the starting time for school, and how the child gets to school; agreement figures for these questions were 75.0%, 56.3%, and 85.4%, respectively. In short, it appears that open-ended questions requesting factual information yield responses of varying validity, depending on the questions. Apparently questions involving number and time concepts were difficult. The discouragingly low agreement on some of these questions encourages a conservative attitude when interpreting data drawn from interviews with mentally retarded respondents.

Open-Ended Questions with Coded Response Categories

Three open-ended questions posed of community children and their parents required subsequent coding of responses into categories. For example, responses to the questions "What things would you really like to learn in school?" were assigned to such categories as academics, physical recreation, and self-care skills. For each of these questions, ten such response categories were defined. Comparisons were conducted on a category-by-category basis, so that each category was coded either "mentioned" or "not mentioned" by client and parent. Responses were then compared in a 2X2 contingency table, computed by averaging across all 30 response categories (Table 7.5). Average agreement of these 30 categories was 77.9%. However, as can be seen, most of this agreement resulted because both child and parent did not mention a category. Only 6.8% of client-parent comparisons yielded agreement because both client and parent mentioned a particular category. It is difficult to draw conclusions from these data, because agreement depends to a large extent

Table 7.5: Mean Contingency Table for 30 Coded Response Categories to 3 Open-ended Questions

Responses of clients	Response of Significant Others		
	Did not mention	Mentioned	
Did not mention	71.1%	10.7%	81.8%
Mentioned	11.4%	6.8%	18.2%
Total	82.5%	17.5%	

Agreement = 77.9%

on the verbosity of respondents and the number of categories developed by the coders. It is at least clear that these open-ended questions yielded very little information. It is also instructive to note that the four response categories which yielded the highest proportion of mention-mention agreements were the four categories lowest in overall child-parent agreement. This suggests that when clients and their significant others do offer information in response to open-ended questions, the responses they make frequently differ. In view of this lack of evidence for the validity of responses to open-ended questions, as well as the low responsiveness rates for these questions, it would appear that this format holds little promise in interviewing mentally retarded persons.

Conclusions

Analysis of the open-ended questions suggests two major conclusions. First, information drawn from open-ended questions must be evaluated with caution. The low agreement obtained on simple factual questions and the large variability in agreement across questions support this observation. Second, the open-ended question format is a relatively unattractive one. These questions tend to pull very little information from mentally retarded persons, and the information obtained from open-ended questions has not been validated by agreement with parents' responses.

CONCLUSIONS

General Considerations

These data regarding the validity of the responses of mentally retarded persons are not encouraging. It is clear that information drawn from interviews with retarded persons must be evaluated critically.

Children's responses frequently fail to correspond closely to the responses of their parents. It is particularly notable that this lack of agreement exists for some simple questions regarding factual information which should be readily available to most persons. Second, it is apparent that different questions yield responses of different validity. In some cases, particular topics seem to be affected by systematic response biases. Furthermore, conceptually more difficult questions appear to yield less valid responses. However, it seems difficult to predict in advance which questions will yield very low client-parent agreement. This fact is of vital importance because it implies that response validity can never be globally assumed; rather, each question must be interpreted conservatively since any individual question may be especially subject to response effects. Unpredictable variability, coupled with the very low validity of many responses, requires that researchers who interview mentally retarded persons design interviews which deal very carefully with the issue of validity.

Format Recommendations

The fact that major differences exist in client-parent agreement as a function of question format suggests that some formats are preferable to others. Yes-no questions appear especially unsuitable due to the effects of acquiescence. Client-parent agreement frequently drops to below chance levels when a high proportion of parents respond "no" to these questions. Either-or questions, on the other hand, appear to have particular promise. Agreement for these questions was relatively high and appeared to be uninfluenced by

systematic biases. Multiple choice questions offering response alternatives which differ quantitatively are very unattractive; agreement on these questions was below chance level in this sample. In contrast, multiple choice questions offering discrete response alternatives and dealing with basic factual information yielded the highest agreement rates of any format. Further exploration of the potentials of discrete multiple choice questions seems appropriate. On the basis of client-parent agreement the open-ended question format appears undesirable. Average agreement was reasonably high when the question called for a single factual answer, but open-ended questions calling for an active listing of responses are probably useless. Moreover, the fact that many retarded persons cannot answer open-ended questions means that even when they yielded valid answers those answers come from only a limited segment of the population. In short, either-or questions or multiple choice questions offering discrete alternatives would appear to have the most promise as methods of obtaining information from mentally retarded respondents.

ANALYSIS OF AGREEMENT USING COEFFICIENT KAPPA

In the preceding discussion of agreement in the community children's sample, little mention was made of any standard against which agreement figures can be compared. We have simply compared agreement figures and applied an intuitive standard based on the anticipated ranges of agreement figures. At one extreme, if responding were completely random on the part of both clients and significant others, agreement for questions with two response alternatives would average 50%. At the other extreme, perfect agreement between clients and their significant others

would result in agreement of 100%. An agreement figure can be roughly evaluated according to where it falls on this continuum. Agreement lower than or little better than 50% suggests low validity, while agreement approaching 100% is impressive. (One must, of course, remember that agreement is a rough estimate of validity; e.g., for some questions, acquiescence inflates agreement but actually raises questions about response validity.) Using this intuitive standard, the 82.0% agreement for the hypothetical questions depicted in Table 7.6 appears to be reasonable evidence of response validity.

However, as was done in Chapter 6 on reliability, kappa coefficients can be calculated to compare obtained agreement to "chance" agreements as defined by reference to the response patterns (i.e., marginals) of the two groups. Given the marginals depicted in Table 7.6, the extent of agreement expected by chance for the hypothetical question is 82.0% ($.9 \times .9 + .1 \times .1$). That is, given that 90% of clients and 90% of significant others respond "yes" to this question, then 82.0% of client-significant other pairs would agree if those pairings were established by drawing randomly from the two populations. Thus, the actual agreement figure of 82.0%, impressive according to the intuitive standard, is actually no different than the extent of agreement to be expected by chance, and the kappa analysis makes us wary of making too much of the extent of agreement obtained.

If we were concerned only with the aggregate level of analysis, reference to agreement figures and associated kappas would not even be necessary. It would be adequate to look only at the marginals of contingency tables. If, as in the

Table 7.6: Hypothetical Yes-No Question

Clients	Significant Others		
	No	Yes	Total
No	1.0%	9.0%	10.0%
Yes	9.0%	81.0%	90.0%
Total	10.0%	90.0%	

Agreement = 82.0%

example in Table 7.6, the percentages saying "yes" in each group were identical or nearly identical, we would have a basis for concluding that the picture of the sample obtained is virtually the same whether one listens to clients or to their significant others. This type of concurrence would, of course, be possible even if there were numerous discrepancies between individual clients and their significant others.

If, on the other hand, one is concerned (as we have been) with validity at the individual level of analysis, attention to paired responses is required. We have generally been content to use an intuitive standard in judging agreement figures, for we feel that 70 to 80% or greater agreement between clients and significant others is impressive. Moreover, it means that we would have confidence, if we were to conduct correlational analyses using either client or significant other data in its place, that findings based on the two groups of informants would generally be similar. However, we can still ask whether a given agreement figure really represents agreement beyond

what one would obtain if client and significant other responses were paired randomly rather than correctly.

It is for this purpose that we have chosen to use kappa selectively simply to provide an additional perspective on consistency and agreement by chance; it is difficult for kappa to be large and significant when the marginals are lopsided simply because the probability of agreement by chance becomes large. Given an agreement figure 80 to 90%, we can still be satisfied that clients and their significant others are saying much the same things, even if kappa is not significant. On the other hand, we cannot be fully satisfied by low agreement figures even if kappa is large, for we would not be confident using data from clients and significant others as if they were interchangeable.

We can be most satisfied if agreement is high and kappa is high, for then we have reason to conclude that the extent of agreement obtained is greater than what would be expected if responses from the two groups were paired randomly.

AGREEMENT IN THE COMMUNITY CHILDREN'S SAMPLE BY THE KAPPA CRITERION

Overall Agreement

The use of kappa to analyze agreement yields discouraging results. Of 88 client-parent comparisons for the community children's sample, only 11 resulted in agreement which differed significantly from chance (at the .05 level) according to this standard.

Agreement by Question Format

Multiple choice questions offering discrete response alternatives emerge as the most desirable format when kappa is the criterion. All six of these multiple choice questions generated agreement significantly different from chance. To cite an example, for the question "Do you live in a duplex, a house, a trailer house, or an apartment building?" agreement was 81.6%, with kappa = .44 (p less than .05, one-tailed). It is unfortunate that these six questions all dealt with either the client's means of transportation to school or with his/her dwelling place, as these topics were not adequately treated in other formats. Therefore, the possibility cannot clearly be ruled out that it is these topics, rather than the question format, which engender unusually valid responses. In short, analysis using kappa seems to recommend discrete multiple choice questions even more strongly than did the previous analysis. However, none of the five multiple choice questions using quantitative response alternatives attained significance according to the kappa criterion. Thus the kappas reinforce the conclusion that these types of multiple choice questions are not associated with high validity of response.

Of the 32 yes-no questions in this sample, only three generated agreement significantly different from

chance according to the kappa criterion. All three of these questions were simple verbal questions having to do with teaching skills. For example, for the question "Is anybody at school teaching you about cooking now?" agreement was 71.1%, with kappa = .42 (p less than .01, one-tailed). The low percentage of yes-no questions for which agreement was significant by this criterion highlights the conclusion based on the earlier analysis: yes-no questions do not elicit generally valid responses from mentally retarded respondents. However, it should be noted that the high proportions of "yes" responses on many of these questions mean that a large kappa is difficult to obtain.

The remaining two comparisons for which kappa indicates agreement differing significantly from chance were open-ended questions. One of the four open-ended factual questions, "Most days, how do you get to school?" resulted in agreement of 85.4% (kappa = .68, p less than .01, one-tailed). The fact that the content of this question coincides with the content of three of the discrete multiple choice questions may suggest that the especially impressive agreement for these questions is at least partly attributable to question content. Additionally, one of the 30 response categories for open-ended questions with coded responses generated agreement significantly different from chance by the kappa criterion. For the question "Do you play any sports? (if yes) What sports do you play?" the category "basketball" generated agreement of 75.0% with kappa = .43 (p less than .01, one-tailed). Given only two significant agreement figures for the 34 client-parent comparisons for the open-

ended questions, the Kappa analysis does not appear to challenge our earlier conclusion that open-ended questions are not particularly appropriate for interviews with mentally retarded respondents. Again, however, the high proportions of "no mention" responses among both children and parents make the probability of agreement by chance correspondingly high.

One conclusion based on the intuitive analysis of agreement is somewhat in opposition to the conclusions drawn from the kappa analysis. That is, none of the eleven either-or questions in this sample resulted in agreement differing significantly from chance by the kappa criterion. Because our earlier analysis suggested that either-or questions are, along with discrete multiple choice questions, the most desirable question type, this observation is problematic. We can again point out that the marginals for either-or questions are lopsided, elevating the standard for significant agreement. We can still conclude that agreement is substantial and that systematic response biases (e.g., selection of the last option) are relatively weak, but we are forced to be more tentative about the merits of either-or questions.

In the analyses of agreement in the remaining samples which follow, we will not present kappa statistics. Our discussions will emphasize the sheer degrees of agreement obtained as well as the systematic biases accounting for client-significant other disagreement. Conclusions based on kappa statistics are generally parallel from sample to sample. Repeatedly we find that simple questions yield similar re-

sponse breakdowns, often lopsided ones, among both clients and significant others. Whether or not agreement exceeds chance expectation is perhaps less important than whether or not clients and significant others are in substantial agreement.

Additional Samples

Extensive comparisons were made of the responses of clients and attendants in both the institutionalized adult and institutionalized children samples. The vast majority of this information echoed the conclusions stated above. The following discussion will examine those instances in which data from the institutionalized children or adults either contradicted or supplemented the conclusions drawn from the community children sample.

Overall Agreement

A total of 126 questions permitted comparison of responses given by institutionalized adults with responses given by their attendants; 56 questions were asked of both institutionalized children and their attendants. Average agreement figures for these samples and for the community children are compared in Table 7.7. The overall agreement figures for the three samples are remarkably similar, suggesting that whether informants are parents or attendants one can generally expect mentally retarded persons to agree with them slightly less than two-thirds of the time. Where questions of one type were asked of more than one sample, the similarity in agreement figures is impressive. For the most part, the conclusions drawn from analysis of agreement in the community children's sample apply equally to the data from additional samples.

Table 7.7: Agreement by Question Type: All Samples

Question type	# of Questions			% Agreement		
	Community Children	Inst. Adults	Inst. Children	Community Children	Inst. Adults	Inst. Children
Yes-No (words)	22	62	35	65.7	60.6	63.0
Yes-No (with pictures)	10	10	13	71.5	72.1	60.1
Either-Or (words only)	7	4	2	65.6	71.4	72.3
Either-Or (pictures)	4			58.1		
Multiple choice (discrete)	6			77.9		
Multiple choice (quantitative)	5	26	1	24.0	33.2	25.6
Open-Ended (factual)	4	4		64.3	81.6	
Open-Ended (numerical)		1			7.6	
Open-Ended (coded categories)	3	19	5	77.9	79.9	72.0
	61	126	56	64.5	59.3	62.8

ADDITIONAL INFORMATION FROM THE INSTITUTIONAL SAMPLES

Yes-No Questions in the Institutionalized Adult Sample

In the discussion of the community children's sample, it was suggested that a group of questions to which 50% of the responses of significant others were "yes" would give an improved estimate of response validity. The simple yes-no questions for the institutionalized adult sample provide just such an estimate. For 22 simple yes-no questions in the community children sample, 68.8% of parents' responses were "yes," and agreement was 65.7%. For the 44 simple yes-no questions asked of institutionalized adults, 47.7% of attendants' responses were "yes," and agreement dropped to 57.7%. It is apparent that when agreement is not inflated by acquiescence, agreement for yes-no questions is only slightly better than the 50% that would be expected if responding by both clients and their significant others were completely random.

Rules Questions

Eight yes-no questions dealing with rules ("Are you allowed to stay up late at night?") were posed of institutionalized adults and yielded a pattern of responses quite different from that typifying yes-no questions in general. For four questions, the wording was "Is it against the rules...?" and for the other four, the wording was "Are you allowed...?" All questions dealt with behaviors that are generally prohibited. These questions were paired so that the two sets of four dealt with precisely the same content but were worded oppositely. Agreement for the "Is it against the rules...?" questions was only 53.4%, with 35.3% of the

comparisons disagreements in which the clients responded "no" and the attendant "yes." Agreement for the "Are you allowed...?" questions averaged 79.9%, with 70.3% of comparisons agreements in which both respondents answered "no." It appears that questions dealing with obvious rules and phrased "Are you allowed...?" are unusually conducive to valid responding. By contrast, respondents appear to have misunderstood the "Is it against the rules...?" questions and may have responded "no" thinking that they were indicating that the behavior was not allowed. Even more interestingly, agreement for these questions was clearly not influenced by acquiescence since most agreements consisted of "no" responses from both client and attendant. Retarded persons do not acquiesce to just any yes-no questions. (In fact, here the correlation between proportion of attendants responding "yes" and agreement was $-.75$, a direct reversal of the more usual trend.) These questions, in combination with the skills questions discussed earlier, suggest that the retarded are most likely to acquiesce when that is the socially desirable response but do not when "no" appears to be socially desirable. Clients appeared to have a clear sense that the activities asked about were socially unacceptable.

Last-Option Position Bias in Either-Or Questions

Within the either-or questions asked of institutionalized adults, there was evidence for a weak position bias on the part of mentally retarded respondents, who showed a slight preference for the last-mentioned alternative. Asked to specify whether they are usually happy or sad, 91.2% of clients responded "happy" when that was the last-mentioned alternative, as

opposed to 84.3% when "sad" was mentioned last. When asked if they are usually alone or with others, 64.5% of respondents said "with others" when that alternative was mentioned last, as opposed to 51.5% when "alone" was mentioned last. This response set tends to elevate agreement slightly when the last-mentioned alternative is the response preferred by attendants. However, this effect is very small by comparison with the effects of acquiescence. Agreement for either-or questions in the adult institutional sample supported our earlier conclusion that the validity of responses to either-or questions exceeds the validity of responses to yes-no questions.

Open-ended Questions with Preset Categories

Four questions asked of institutionalized adults and their attendants were open-ended questions for which responses were coded into one of two preset categories. For example, all responses to the question "Who decides how you spend your money?" were coded into the categories "somebody else" or "client." Average agreement for these questions was 81.6%, ranging from 70.6% to 92.6%. (A fifth open-ended question "How many people sleep in your bedroom?" required numerical responses and resulted in agreement of only 7.6%. Interviewees were totally unable to provide valid answers to this question.) The average agreement figure of 81.6% for the four questions far exceeds that of 64.3% for the open-ended factual questions in the community children's sample. It appears that open-ended questions soliciting simple information and using preset categories may be more promising than was concluded on the basis of interviews of community children. In fact, the

unusually high agreement for these questions casts this format in a very favorable light, although it is still the case that many interviewees cannot answer such questions.

Conclusions

In short, the data from additional samples modify the conclusions drawn from the community children sample in a few ways. First, acquiescence does not occur to all questions. The "rules" questions showed that under specific circumstances, this response bias can be outweighed by other factors (possibly a social desirability response set). Second, either-or questions in some cases engender a last option position bias, which to some degree reduces the validity of responses to these questions. Third, open-ended questions with preset response categories may be somewhat more useful than was concluded on the basis of responses of community children. In general the conclusions drawn from the community children sample are supported by these further analyses.

Correct-Incorrect Questions as a Measure of Response Validity

ALL QUESTIONS

Although client-significant other agreement served as our basic measure of response validity, a number of questions posed of mentally retarded respondents had known correct answers, allowing for another type of analysis of validity. Of nineteen such questions, the responses to twelve were coded as either incorrect, partially correct, or correct, while responses to the remaining seven were coded either incorrect or correct. For the entire group of twenty questions, 77.1% of clients' responses were correct or partially correct. For the twelve incorrect-partially-correct

questions, an average of 18.9% of responses were incorrect, 19.9% were partially correct, and 61.2% were correct. For the seven questions coded simply incorrect or correct, an average of 29.7% of responses were incorrect, and an average of 70.3% of responses were correct. Because such average figures are difficult to conceptualize, these questions are discussed on a more detailed level below.

QUESTIONS ASKED OF ALL SAMPLES

Two open-ended questions were asked in identical form of respondents in the three major samples included in our study: community children, institutional adults, and institutional children. The first question, "What is your full name?" yielded the results depicted in Table 7.8. Answers were coded as partially correct if the respondent made an obvious attempt to say either the first or last name, and correct if both the first and last names were stated correctly. In general, the data in Table 7.8

reveal that about three-quarters of clients who can answer the question are able to give their full names correctly, and another quarter are able to state either their first or last name. Only one of the 108 clients who responded appropriately to this question was unable to give at least a partially correct answer. In one sense, this seems somewhat encouraging. However, it is important to remember that many clients were unable to respond to this question. In fact, when those who were asked but were unresponsive to this question are considered, only 64.5% of all community children, 40.2% of institutionalized adults, and 50.0% of institutionalized children who were asked the question were able to correctly state both first and last names. Since one's name is in a sense the most basic piece of information that anyone possesses, consideration of this question could perhaps serve as a sobering foundation for our thinking regarding interviewing retarded persons. If only around half of mentally retarded persons can correctly state their

Table 7.8: Percentage of Correct Responses to "What is your full name?"

	Incorrect	Partially Correct	Correct
Community children	0.0%	22.9%	77.1%
Institutionalized adults	3.3%	30.0%	66.7%
Institutionalized children	0.0%	26.7%	73.7%

full names in response to a direct question, then the implication of the previous portion of this chapter, that the validity of the responses of mentally retarded individuals is frequently minimal and always in question, is no way surprising. The most optimistic interpretation of this information is that when retarded persons do respond appropriately to very simple questions, their responses can be expected to be at least partially correct.

Additionally, the question "How do you spell your name?" was asked of the three samples. In response, 74.5% of community children, 71.4% of institutionalized adults, and 82.8% of institutionalized children have responses which were at least partially correct. This information perhaps pertains more to our respondents' language skills than to response validity per se.

QUESTIONS ASKED OF TWO SAMPLES

Two additional questions were included in the interviews with both community children and institutionalized adults. The first of these questions, "Please write your name for me on this paper," was again more a test of language skills than of response validity per se.

Among community children, 77.8% of respondents gave correct or partially correct answers, and 60.0% of institutionalized adults were at least partially correct.

The second question was phrased slightly differently for the two samples. "What is your address at home?" was the wording for the community children, as opposed to "What is your address here?" asked of institutionalized adults. Responses of institutionalized clients were coded correct if they mentioned both the institution and the city, and partially correct if they mentioned one but not the other. Responses of community children were coded correct if they gave a mailing or street address, and partially correct if they gave a street name without a house number. The fraction of correct responses, depicted in Table 7.9, gives us another solid insight regarding the issue of response validity. This question involves information which should be readily available to most persons, but it was apparently considerably more difficult than the "name" questions.

ADDITIONAL QUESTIONS

A number of additional factual questions were asked of only one sample. Three open-ended questions,

Table 7.9: Percentage of Respondents Correctly Stating Their Addresses

	Incorrect	Partially Correct	Correct
Community children	21.4%	14.3%	64.3%
Institutionalized adults	25.0%	16.7%	58.3%

coded incorrect or correct, were asked of institutionalized adults. These three questions, "What month is your birthday?" "What day is your birthday?" and "What year were you born?" resulted in 80.6%, 77.8%, and 64.7% correct responses, respectively. Again, those clients who were able to respond appropriately generally gave accurate information. However, if we take into account unresponsive persons, only 28.2% of those clients asked the question were able to state their day of birth, 48.6% correctly named the month in which they were born, and only 20.1% correctly named their birth year. In short, mentally retarded respondents are often unable to correctly answer even very simple questions, but this inability usually takes the form of a failure to respond appropriately, so that those who do respond appropriately are likely to also respond correctly.

Four yes-no questions having known correct answers, all involving the value of money, were included in the interview of institutionalized adults. These questions were included to assess subjects' understandings of money concepts. For the questions "Could you buy some candy with \$100?" "Could you buy a new bicycle with \$1.00?" "Could you buy a new TV with \$10.00?" and "Could you buy a record album with \$10.00?" the fraction of correct responses was 38.9%, 60.5%, 27.0%, and 92.3%, respectively. It is especially interesting to note that the questions for which the correct answer was "yes" had by far the highest fraction of correct responses, another instance of the effect of acquiescence. This factor makes it difficult to evaluate the validity of responses to these questions. The fact that 50% of respondents would be expected to answer these questions correctly by chance and only 27.0% correctly answered the TV question casts serious doubt on whether re-

tarded persons can be expected to answer questions about the value of money.

Finally, two open-ended questions were asked only of the institutional children. In response to the question "What is your birthday?" 66.7% of children gave correct responses, 27.3% gave partially correct responses and 6.1% gave incorrect responses. For the question "What kind of place is this?" 61.9% responded correctly (by referring to the institution) and 38.1% incorrectly. Once again, our clients showed a clear but limited ability to respond correctly to simple questions, if they can respond at all.

CONCLUSIONS

As with the agreement data, the results of these factual questions must leave us pessimistic but not without hope. Mentally retarded individuals do tend to respond correctly to these simple, straightforward factual questions. However, a discouragingly high proportion of respondents respond incorrectly when very fundamental information is requested of them: name, birth date, address, and so on. Furthermore, when responsiveness data are considered in tandem with information about validity, it is clear that only a small proportion of retarded persons can correctly answer even these "easy" questions. Examples of some of the difficulties we encountered are included in Sample 7.1. These correct-incorrect questions provide a succinct summary of all our data regarding validity, in that they demonstrate that mentally retarded persons can, under some circumstances, provide valid information about themselves, but that the validity of the information they can provide can never be taken for granted.

Sample 7.1

A Sampler of Problems in Obtaining Basic Factual Information

1. Q: What is your full name?
A: (Lifted her foot and pointed to the bottom of shoe, where her name was written)
2. Q: What is your birthday?
A: I don't know. They got it in the records. I'm 15. I think it's February.
3. Q: What is your name?
A: My name is Ronnie Peters.
Q: How do you spell your name?
A: RONKLTUWXYZ
4. Q: What year were you born?
A: Little Rock.
5. Q: Do you get SSI?
A: What?
Q: Do you get SSI?
A: SSI? Spell it!!
6. Q: What is your address here?
A: 35944234
7. Q: What is your full name?
A: Huh? I don't got no name.
Q: Do you get SSI?
A: I gots my bus ticket.
Q: (repeated)
A: I guess.
Q: Do you have a family?
A: No except my mama and daddy. That's all I have.
Q: (repeated)
A: No.

Agreement With Parents and Attendants as an Individual Characteristic

To provide another perspective on client-significant other agreement, we focused on 13 yes-no questions that were asked in precisely the same form to institution children, community children, and institution adults. Nine of the questions were about par-

ticipation in various chores and were accompanied by pictures, while the remaining four were simple verbal yes-no questions. An agreement score was calculated for each subject who had responses paired with significant other responses on at least half of the set of 13 questions. The agreement score could range from 0 to 100%. We sought to determine whether agreement differed from sample to sample

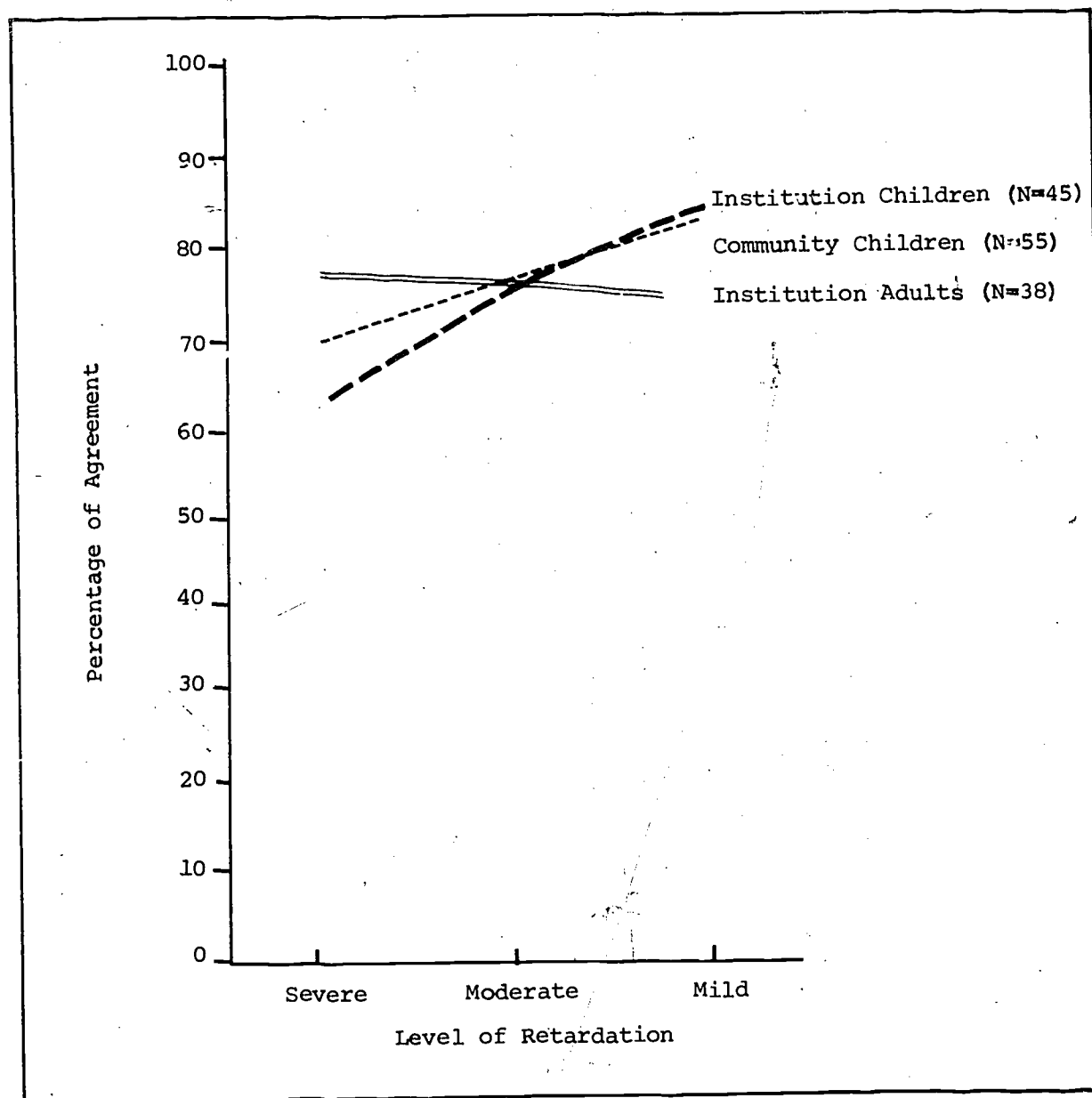


Figure 7.2 Agreement with Parent By IQ Group for 13 Yes-No Questions Used in Three Samples

when the same questions were at issue and to explore individual differences in the tendency to agree with parents or attendants.

Figure 7.2 presents the mean agreement scores for severely, moderately, and mildly retarded subjects in each of the three samples. As the figure indicates, there was some tendency for agreement with significant other to increase as IQ increased. A 3 X 3 analysis of variance indicated that there was a significant difference between IQ groups ($F(2, 129) = 3.62$, $p = .03$), but that the three samples were not significantly different from each other and that IQ group and sample did not interact. Overall, severely retarded persons agreed with their significant others 70.5% of the time, moderately retarded persons 74.4% of the time, and mildly retarded persons 79.7% of the time on these questions. Follow-up t-tests with the three samples collapsed indicated that severely retarded persons were significantly less likely to agree with their significant others than were mildly retarded persons, $t(135) = 2.56$, $p = .012$. Although the interaction between IQ group and sample was not significant, there were some variations in findings from sample to sample. Among institutionalized children, the severely retarded group was definitely deficient in comparison to the mildly retarded group, $t(42) = 2.66$, $p = .02$. In the community children's sample, the difference between severely retarded and mildly retarded respondents approached significance ($p = .056$). Interestingly, there were no significant IQ group differences in the adult institution sample, and in fact agreement scores actually decreased slightly as IQ increased. Altogether, then, this analysis revealed a tendency for agreement with significant other to be a func-

tion of level of retardation, although it was certainly not a strong pattern even though it was statistically significant. Moreover, level of agreement did not differ greatly from sample to sample, despite the fact that the parents of children living in the community might be expected to be more knowledgeable informants than institution attendants responsible for many residents. Of course, it must be borne in mind that this analysis could only be conducted on yes-no questions, which, as discussed previously, often appear to yield biased answers which may inflate estimates of validity as indicated by agreement with significant others. Nonetheless, for such questions one can apparently expect agreement with significant others, whether they are parents or institution staff, to average about 75% and to be weakly but positively related to the respondent's IQ.

To show more concretely the levels of agreement which were obtained, Tables 7.11, 7.12, and 7.13 report for each sample the percentage of clients and significant others saying "yes" to each question and the percentage of cases in which answers agreed. As noted in Table 7.12, there was only one instance in which the kappa statistic indicated agreement beyond the level expected by chance (cleaning floors in the adult institution sample). At the same time, the percentages agreeing are exceptionally high on many other questions, and it must be borne in mind that extremely lopsided distributions in which virtually everyone says "yes" mitigate against obtaining statistically significant agreement. The tables add still another perspective, for in some cases, despite relatively low agreement figures, the proportions of clients and significant others saying "yes" are very similar. For

Table 7.11: Agreement in Children's Institution Sample

Activity	% Yes according to clients	% Yes according to SO	Percent Agreeing	Kappa
Watch TV	97.8	97.8	95.6	-.022
Listen radio, record player	92.9	100.	92.9	.0217
Go to church	89.1	100.	89.1	.000
Count money (knows how)	88.5	42.3	38.4	.101
Set table	75.6	95.6	75.5	.081
Do dishes	90.7	58.1	62.8	.145
Clean the floor	71.1	95.6	71.1	.060
Dust furniture	83.7	88.4	76.7	.034
Make bed	90.7	90.7	86.0	.170
Pick up stuff	93.2	95.5	88.6	-.039
Cook on the stove	61.0	7.3	41.5	.015
Make sandwiches (food)	89.2	24.3	29.7	-.002
Take out trash	97.7	79.1	81.4	.164

Table 7.12: Agreement in Community Children's Sample

Activity	% Yes according to clients	% Yes according to SO	Percent Agreeing	Kappa
Watch TV	98.2	100.0	98.2	.000
Listen radio, record player	90.9	100.0	90.9	.000
Go to church	81.8	85.5	78.4	.212
Count money (knows how)	80.8	61.5	53.9	-.074
Set table	74.5	76.4	72.7	.263
Do dishes	70.9	72.7	69.1	.237
Clean the floor	76.4	74.5	65.5	.069
Dust furniture	76.4	78.2	65.5	.017
Make bed	85.7	82.1	75.0	.077
Pick up stuff	89.1	96.4	85.5	-.057
Cook on the stove	65.5	52.7	47.2	-.074
Make sandwiches (food)	89.1	90.9	83.6	.090
Take out trash	83.6	89.1	80.0	.157

Table 7.13: Agreement in Adult Institution Sample

Activity	% Yes according to clients	% Yes according to SO	Percent Agreeing	Kappa
Watch TV	97.5	97.5	95.0	-.026
Listen radio, record player	89.7	100.0	89.7	.000
Go to church	94.7	100.0	94.7	.000
Count money (knows how)	71.9	31.3	53.1	.194
Set table	79.5	76.9	61.6	-.125
Do dishes	78.9	78.9	73.7	.210
Clean the floor	78.9	81.6	81.6	.420*
Dust furniture	84.2	92.1	81.5	.128
Make bed	89.2	94.6	83.8	-.078
Pick up stuff	83.8	94.6	78.4	-.088
Cook on the stove	67.6	8.8	35.3	-.003
Make sandwiches (food)	75.0	69.4	66.7	.174
Take out trash	86.5	97.3	83.8	-.047

* $Z = 1.77$, significant at the .05 level.

example, while institutionalized adults and their attendants were in agreement only 61.1% of the time about the client's involvement in setting the table, the 79.5% of the clients indicating involvement is not very different from the 76.9% of the attendants indicating involvement. In such cases, the implications of disagreement between clients and significant others are not severe if one is only interested in aggregate figures, for despite disagreement at the individual level, pictures of the group emerging from the two groups of informants are much alike. In other cases, however, it is clear that one does not obtain the same information from the two groups of informants. The most markedly different pictures of the group were obtained on the questions about whether the client knows how to count money and cooks on the stove. According to the clients themselves, for example, 88.5% of the institution children, 80.8% of the community children, and 71.9% of the institution adults know how to count money. The corresponding figures based on significant other responses were only 42.3%, 61.5%, and 31.3%, respectively. Similarly, according to clients 61.0% of the institution children, 65.5% of the community children, and 67.6% of the institution adults cook on the stove. The quite different figures provided by significant others were 7.3%, 53.7%, and 8.8%, respectively. In both cases, clients, by saying "yes" frequently, are apparently overstating their involvement. These figures reinforce once more the variations in agreement associated with individual content areas and the value of interpreting agreement from more than one perspective.

CORRELATES OF AGREEMENT

Given the fact that acquiescence on yes-no questions has emerged as

a problem in interviewing the mentally retarded, we attempted to look at agreement with significant others in relationship to acquiescence. We correlated the agreement score described above with a dichotomous acquiescence score which simply indicated whether or not a client responded "yes" to both the question, "Are you usually happy?" and the question, "Are you usually sad?" To further explore the nature of agreement with significant other as an individual characteristic, we also correlated the agreement measure with IQ, sex, and a responsiveness score developed on the basis of 22 items used in all three samples. Table 7.14 presents these correlations for each of three samples.

What does this table reveal about agreement? First, as was indicated in the previous discussion, higher IQ respondents were more likely to agree with their significant others except in the sample of institutionalized adults. Moreover, those clients who were able to answer a greater proportion of questions appropriately were generally more in agreement with their significant others, although again this relationship does not hold in the adult sample. (It should be noted that agreement scores were calculated only if a client had been able to respond to at least half the questions in the set of 13, and the correlation coefficients are based only on those subjects who had scores.) In the children's institution sample only, males were more likely to agree with significant others than were females, while in the community children's sample this relationship was reversed in direction, and in the adult sample nonexistent, leaving us uncertain about the relationship between sex and agreement. Finally, in the children's institution sample, respondents who acquiesced were significantly less

Table 7.14: Correlations of Measures in Three Samples

	IQ	Sex	Agree	Response	Acquiescence
IQ	Inst. child	-.13	.40**	.63**	-.40**
	Comm. child	-.02	.33**	.33**	-.30*
	Instit. adult	.03	.02	.31*	-.31*
Sex			-.28*	-.02	-.02
			.38**	.11	-.06
			-.08	.01	-.14
Agreement				.37**	-.32*
				.20	-.13
				.07	-.03
Responsiveness					-.44**
					-.21
					-.38**

*Significant at the .05 level

**Significant at the .01 level

Note: In each cell of the matrix, correlations in order moving downward are for the institution children, community, and institution adult samples, respectively. Sex is coded male = 1, female = 2.

likely to agree with their significant others, but this negative relationship was weak in the other samples. The relationship between acquiescence and agreement is difficult to unravel. Lower IQ clients, while they tend to disagree with informants, also tend to acquiesce a great deal, and as we have noted previously acquiescence may inflate agreement if informants often say "yes." Curiously, relationships were consistently stronger in predicted directions in the children's

institution sample than in other samples. Generally, trends in the same direction were evident among community children, but none of the relationships held among institutionalized adults. As this table also indicates, IQ was quite consistently related to responsiveness in a positive direction and to acquiescence in a negative direction (i.e., higher IQ subjects were more responsive and less likely to acquiesce), and similarly, highly responsive subjects were less likely to

acquiesce than were less responsive subjects.

Conclusions

This chapter has examined the validity of answers given by mentally retarded interviewees in two ways; by comparing their answers with parallel responses given by attendants or parents, and by comparing their answers with known fact. As this analysis has indicated, establishing whether or not answers are valid, especially through the first approach, is problematic. We had no alternative most of the time but to look at each question individually and then try to detect patterns in the data for sets of questions.

What can we conclude overall? We can conclude, first of all, that the information provided by retarded persons and the information provided by non-retarded informants about retarded persons is not necessarily the same. Indeed, in many cases one would obtain very different pictures of the needs, attitudes, and circumstances of retarded citizens from the two types of respondents. Our data suggest that their answers match approximately two-thirds of the time overall. The implications of this basic finding are serious. Mental retardation researchers (and practitioners as well, in their informal information-gathering efforts) have not been sensitive to this phenomenon. They have often relied on one or other source of information, assuming, but only assuming, that data from the two sources would be interchangeable. To give an example, consider the literature on deinstitutionalization, which includes studies

ed on client interviews and studies based on interviews with institution staff, houseparents, and other nonretarded informants. Studies based on different sources of data can be compared, but in view of the analyses presented here, one

must recognize that findings may differ from study to study because data were collected from different informants.

At the outset, we warned that one cannot always assume that retarded persons are wrong when they disagree with their parents or attendants. We did indeed find instances in which retarded persons and their significant others simply had legitimate differences of opinion. However, there were other instances in which we would point to response bias on the part of retarded persons as the cause of disagreement. It is these latter cases that allow us to conclude that retarded persons are sometimes "wrong" in the sense that their answers do not reflect the state of things as much as they reflect the operation of response sets. The clearest case of all is the threat to validity of response to yes-no questions represented by acquiescence. As a general rule, retarded persons and their parents or attendants agreed on yes-no questions only to the extent that parents and attendants happened to say "yes" frequently, thus matching the frequent "yes" responses of clients. This phenomenon raised serious questions about the very use of agreement figures as a measure of validity of response. In short, we had to take into account the fact that high agreement was not necessarily evidence of validity, not when we could identify other factors such as acquiescence that were artificially inflating agreement figures.

We encountered similar problems in analyzing other types of questions. For examples, since very few responses were obtained to open-ended questions calling for enumerations, high agreement figures for categories of response to such questions generally said little more than that clients and nonretarded infor-

ments frequently did not mention something. Thus it must be recognized that many of the average agreement figures presented in this chapter are misleading indicators of validity of response. While we have generally found that retarded people and their significant others agree about two-thirds of the time, that does not mean that responses are valid about two-thirds of the time.

On the positive side, we did find either-or questions to have respectably high agreement figures without the pronounced response bias that characterized yes-no questions. We also found high validity of response to multiple choice questions about basic factual information such as type of residential setting where respondents chose among discrete response alternatives (compared with extremely low validity for multiple choice questions with quantitative response options). And finally we found relatively high validity for factual open-ended questions that required a single answer rather than requiring enumerations of responses. At the same time, none of the agreement figures was high enough that we can be completely confident in the information obtained from retarded persons. We must still conclude that their answers cannot be assumed to be valid unless there is evidence to support that assumption. The inability of many retarded persons to provide basic factual information such as full name, birth date, and address simply reinforced that conclusion.

Is it possible to predict which retarded persons are likely to provide valid answers? We attempted to answer this question by calculating agreement scores for individuals reflecting their tendency to agree with their significant others on a set of yes-no questions that were used in all three major samples. We found that children living in the community were, surprisingly, no more likely to agree with their parents than institutionalized persons were to agree with their attendants, even though we had expected parents and children to be more likely to share a common set of facts and perspectives. While we found that higher IQ persons were more likely than lower IQ persons to agree with significant others, the relationship was not really strong enough that one could use it to predict in advance whose answers will be valid. This matter needs to be explored further, since our findings, which were restricted to yes-no questions, might not generalize to other types of questions, particularly in view of the acquiescence problem for yes-no questions. Until further research is done, one must be wary of responses given by any retarded person, especially on the kinds of questions that we have found to elicit biased responses.

In the next chapter, we will pursue the topics of response bias and response validity through systematic comparisons of alternative questioning approaches.

WHAT INTERVIEWING APPROACHES ARE MOST USEFUL?

A major focus of this report has been evaluation of different kinds of questions according to the criteria of responsiveness, reliability and validity as indicated by agreement with parents and attendants. To facilitate this evaluation, all of the interview schedules used in the various studies included specific comparisons of alternative ways of soliciting the same information, head-to-head competitions among different formats and phrasings applied to the same content area. These comparisons were done specifically to identify questioning techniques that optimize responsiveness, reliability, and validity.

To illustrate the approach used, consider the comparison summarized in Sample 8.1. In comparing two questions, in this case a yes-no question and a parallel open-ended question, we first apply a criterion of responsiveness. We would generally prefer a question that

can be answered by many persons to one that relatively few can answer. In the example, the institutionalized adults studied found it considerably easier to answer the yes-no question than to answer the open-ended question, a finding which is entirely consistent with our analyses of responsiveness as a function of question format (see Chapter 6).

A second criterion centers on reliability in the sense of consistency between answers to the two alternative questions. The consistency of responses to these two questions was relatively low. While that alone does not give us a basis for preferring one question over the other, it is also apparent that there is a systematic trend in the inconsistency. Namely, clients are highly likely to say, "Yes, I decide," in response to the yes-no question while they are mentioning someone other than themselves in response to the open-ended question.

Sample 8.1

COMPARISON OF ALTERNATIVE FORMATS

Yes-no versus Open-ended

(Do you decide what chores you do? versus Who decides
what chores you do? -- Adult Institution Sample)

1. Which Question Yields Higher Responsiveness?

Percent Responding Appropriately:	Yes-no	54.4%
	Open-ended	34.5%

2. To What Extent Are Answers Consistent?

With answers coded client decides versus someone else
decides, percent consistent = 56.3%. All of the rest
(43.7%) are instances of "Yes, I decide" on yes-no and
mention of someone else on open-ended--i.e., acquiescence.

3. To What Extent Do Clients And Significant Other Agree?

For yes-no, 34.6% agree. Remaining 65.4% are cases in
which client says, "Yes, I decide," while attendant men-
tions someone else in response to open-ended question.

For open-ended, 88.2% agree.

CONCLUSION: While more subjects can answer the yes-no question,
acquiescence invalidates response. Open-ended ques-
tion is preferred on grounds of validity.

This suggests that acquiescence in response to the yes-no question is operating and is inflating the number of respondents who claim they make their own decisions.

Finally, we can apply a criterion of agreement with significant others (parents or attendants) in deciding which question is the best approach to use. Due to the acquiescence detected in the consistency analysis, agreement with attendants is lower for the yes-no question than for the open-ended alternative. On the basis of the agreement criterion, the open-ended format is therefore preferable.

As is true in many of the analyses to follow, the application of these three criteria often involves trade-offs. While the yes-no question is to be preferred on grounds of responsiveness, it is clearly deficient on grounds of validity because of the strong influence of acquiescence. Since there is no point in getting answers if they are not trustworthy as valid answers, we must ultimately prefer the open-ended question in this case. Yet neither alternative is entirely attractive and use of the open-ended question will ultimately mean that data will be obtainable only from the more verbal segment of the population under study.

The comparisons which follow are often more complex than the example, but they follow the same general logic. While it will prove impossible on some occasions to clearly establish one alternative as superior to another, we will constantly be trying to determine which ways of asking questions appear to work best in hopes of establishing empirically based recommendations for persons interested in interviewing retarded individuals for research or program planning purposes.

The chapter will consider in turn three major topics in question design. First, we will report on several comparisons of questions whose wording differed but whose formats were the same, asking the question: To what extent are the responses given influenced by the manner in which a question is worded? Secondly, we will briefly consider a procedural issue relevant to the use of open-ended questions: What are the relative merits of simply asking the question as opposed to asking the question and then persistently probing with "What else?" in hopes of eliciting more information? This question is significant given our difficulties eliciting much information in response to open-ended questions.

Finally, we address perhaps the major question underlying our studies: Which question formats appear to be associated with the greatest responsiveness, reliability and validity of response? We began our studies predicting that we would have difficulty getting answers to open-ended questions and that more structured questioning approaches might be more effective. In view of the limited verbal skills of many retarded persons, we also sought to determine whether the use of pictures in interviewing might facilitate answering. We will begin describing our comparisons of alternative formats by considering the use of the multiple choice format in gathering information about activities. Next, we will consider yes-no questions, comparing them with open-ended and either-or formats, as well as assessing the effects of pictures as a means of clarifying yes-no questions. Finally, we will analyze responses to either-or questions and explore the potential value of using pictures as response alternatives in this format.

Most comparisons of alternative formats were conducted a number of times with different content areas. Because of this repetition, it would be unwieldy to present all data comparing two formats. Thus, we will follow the general strategy of selecting one comparison considered to be representative of the others, and presenting that comparison in some detail. Other comparisons of the same type will be cited briefly to supplement information from the illustrative comparison. When results are contradictory or inconsistent, we will attempt to represent that fact and explore reasons for it.

Question Wording

Does the way in which a question is worded affect the ability of mentally retarded persons to answer and the kinds of answers they give? While we devoted more attention to the effects of question format, we did compare a few alternative wordings of questions posed in the same format. (In the alternative formats section we will be discussing several more instances in which, as part of a broader comparison, oppositely worded questions with the same format were evaluated).

USE OF EXAMPLES

Consider first the strategy of including in a question examples of the types of responses one is soliciting. The use of examples (e.g., baseball or football as examples of the class of sports) might be expected to aid mentally retarded persons in understanding a question by making the question more concrete. Alternatively, such examples might bias respondents, leading them to name the items used as examples more than they ordinarily would.

Illustrative Comparison

Two questions asked institutionalized adults to report if they played any indoor games, and if so, which games. One question, "Do you play any games indoors, like cards or checkers? (if yes) Which ones?" gave examples of indoor games. The other, "Do you play any games indoors? (if yes) Which ones?" solicited identical information but did not use examples. Analysis will be limited to the effect of examples on responses to the open-ended portion of each question.

The use of examples increased responsiveness to the open-ended follow-up question. The percentage responding appropriately was 64.5% for the question with examples, as opposed to 52.0% for the question without examples.

The use of examples also influenced the games mentioned in response to the open-ended question. Specifically, only 22.2% of nine adults responding to both open-ended questions mentioned checkers after the plain yes-no question, while 66.7% mentioned checkers after the question in which checkers was included as an example, yielding a consistency figure for this category of only 55.5%. Similarly, respondents were more likely to mention cards in response to the question including cards as an example (55.6%) than to the plain question (22.2%), producing a consistency figure of only 44.4%. Consistency for the remaining categories was generally high and was achieved primarily by not mentioning a game both times. In short, respondents were far more likely to mention a category if that category was cited as an example.

Overall agreement between client and attendant across the seven categories of response to the open-

ended question was 75.5% for the plain question and 72.4% for the question with examples. However, percentages of cases in which both mentioned a category averaged only 10.2% and 4.8%, respectively, with most categories involving no such instances. For the critical categories used as examples, findings were inconsistent. For the checkers category, the use of examples apparently decreased agreement by inducing mentally retarded respondents to overclaim checkers; agreement was 85.7% for the plain question and 60.0% for the question with examples. In the latter case, 40.0% of client-attendant pairings involved the client mentioning checkers while the attendant did not. This trend was not apparent for the cards category; agreement averaged 42.9% after the plain yes-no question and 53.3% after the question with examples, and inconsistency was mixed rather than biased in one direction.

Additional Comparisons

Two other pairs of questions asked of institutionalized adults examined the effects of examples. One pair asked about the client's participation in sports, using baseball and football as examples in one phrasing, and another asked about arts and crafts, mentioning ceramics and painting as examples. The effect of the use of examples on responsiveness to the open-ended follow-up is unclear. For the crafts question, examples apparently increased responsiveness from 41.7% to 59.3%. Conversely, the sports question without examples generated greater responsiveness (75.0%) than the parallel question with examples (57.7%).

These additional comparisons did support the finding that citing an example increases the frequency with which that activity is men-

tioned by mentally retarded persons. For the four examples used in these two pairs of questions, 56.1% of respondents, on the average, mentioned the activity when it was included as an example, versus 44.3% who mentioned the activity when no examples were given. The pattern of inconsistencies generally reflected this bias toward overreporting an activity used as example.

Of these two pairs of questions, only the sports questions were asked of attendants. For the football category, agreement was 55.6% for the question with examples and 57.1% for the plain question. In both cases, most disagreements occurred because the client mentioned football and the attendant did not. Results for the baseball category were far more striking. Agreement was only 11.1% for the question with examples, and 77.8% of comparisons were inconsistencies in which the client mentioned baseball and the attendant did not. For the plain question, agreement was far higher (71.4%). Thus, as for the indoor games questions, the use of examples decreased agreement for one of the activities used as an example but not for the other.

Conclusions

The use of examples in the introductory question had important effects on clients' responses. Although the effect on responsiveness to the open-ended follow-up question is unclear, examples did increase the likelihood that respondents would mention the example activities. The existence of this biasing effect is clear from the consistency data, and client-attendant agreement figures suggest that the use of examples sometimes induces overclaiming by mentally retarded persons, reducing response validity. It appears that the use

of examples with open-ended questions only added to the problems inherent in the basic questioning technique. Examples appeared to arm interviewees with ready answers that they otherwise would not have had.

USE OF ALTERNATIVE TERMS

We also examined the issue of question wording by testing the extent to which the language used to phrase a question affects responses. This strategy was applied with four pairs of questions about rules; one member of each pair asked "Are you allowed to...?" naming a specific behavior, while the other member asked "Is it against the rules to...?" naming the same behavior.

Illustrative Comparison

Two such questions, asked of institutionalized adults, were "Are you allowed to hit people?" and "Is it against the rules here to hit people?" There was no evidence that this difference in question wording altered responsiveness, which was 70.7% for the "against" wording and 70.2% for the "allowed" wording.

Overall consistency for these questions was 43.6%, a figure lower than that expected if responses to both questions were random.

The great majority of respondents (87.2%) answered the "allowed" question "no," while responses were almost evenly split between "yes" and "no" in the case of the "against" wording. A very large fraction of clients, 46.2%, thus contradicted themselves by responding "no" to both questions. This is in startling contrast to the more common result, where acquiescence on oppositely worded questions causes a large proportion of "yes-yes" self contradictions,

and it suggests that acquiescence cannot always be expected on yes-no questions. The most probable explanation is that in the case of the "allowed" questions, "no" was not only the correct response (since hitting people is clearly prohibited), but was also a highly ingrained socially desirable response. This response pattern was carried over to the "against" questions, probably because clients misunderstood the question and believed that "no" indicated that hitting people was forbidden. This interpretation, admittedly speculative, accounts for the differences in the marginals and the preponderance of "no-no" self contradictions.

For the "against" wording, agreement with significant others was 51.3%. Nearly all agreement resulted because the client said "yes" (hitting people is against the rules) and the attendant concurred. Almost all disagreements (43.6% of comparisons) involved a "no" response by the client when the attendant indicated that hitting was forbidden. This again supports the idea that many clients misunderstood the question and attempted, through "no" responses, to give a socially desirable response.

For the "allowed" wording, agreement was 76.3%, with all agreement occurring because client and attendant both said "no" (not allowed). Again, the major conclusion is that these questions, in direct contrast with the remainder of the yes-no questions included in our interviews, did not give rise to acquiescence, perhaps because clients had internalized strong negative evaluations of hitting.

Additional Comparisons

Three additional pairs of questions posed of institutionalized adults pitted "Is it against ...?"

questions against "Are you allowed...?" questions. The activities involved in these questions were calling people ugly names, leaving without asking, and staying up late at night. Response patterns to the first two questions were virtually identical to those for the "hit people" questions, but responses to the last pair of questions differed greatly from responses to the other three.

Mean responsiveness for the three additional "against" questions was 68.4%, compared to 66.1% for the three "allowed" questions. This supports the conclusion that opposite wording did not alter responsiveness. Consistency figures for the "ugly names" and "leave without asking" questions were 40.0% and 57.2%, respectively, with inconsistency predominantly due to responding "no" to both forms of the question. Consistency for the "stay up late" question was 38.2%, with 23.5% of comparison inconsistencies due to saying "yes" both times, and 38.2% inconsistencies due to saying "no" both times. Clearly, staying up late is less strongly prohibited than the other behaviors, leading to far more responses by clients indicating that staying up late is allowed. It is interesting that this factor reduced consistency.

For the three additional "against" questions, average agreement with attendants was 54.1%. Agreement was far higher, averaging 80.8%, for the "allowed" wording. In the latter case, clients' "no" responses coincided with the preferred response of attendants. For the "against" questions, clients' attempts to indicate prohibition through "no" responses operated in opposition to attendants' responses, reducing agreement. This is supported by the observation that most disagreements for the "against"

question (an average of 32.5% of client-attendant pairings) involved a "no" response from the client and a "yes" response by the attendant.

Conclusions

Both consistency and agreement for these questions can be easily interpreted as arising from a strong tendency to say "no" on the part of clients. Why does the usual yea-saying not appear? The simplest explanation is that there is a strong motive to give socially desirable answers; that is, it is highly desirable to express the understanding that prohibited behaviors are prohibited. This accounts for the high proportion of "no" responses to the "allowed" questions. Respondents appeared to understand the rules well enough to give a negative response to negative behavior. Whether they understood the "allowed" phrasing is less clear. One must assume that many clients misunderstood the "against" questions, believing that a "no" response was again the way to express disapproval of undesirable behavior. This interpretation at least raises the possibility that much of the acquiescence observed throughout our interviews may be partially accounted for by social desirability, for questions were typically worded so that "yes" was the socially desirable response. It may be that a tendency toward agreement with socially accepted views (see discussion of Rosen et al., 1977) is in fact the general response set in mentally retarded persons, and that acquiescence to most yes-no questions and nay-saying to the rules questions are simply specific manifestations of that broader tendency. Alternatively, respondents may have been so well-versed on the prohibitions asked about in these questions that their usual tendencies toward acquiescence were overcome.

Regardless of the factors underlying clients' responses, one major conclusion emerges clearly from these comparisons: the picture of a group of mentally retarded persons which is derived from interviews differs as a function of question wording. For the four "against" questions, an average of 61.2% of clients indicated that the behaviors are prohibited. For the four "allowed" questions, an average of 75.4% of clients indicated that the behaviors are prohibited. In neither case do clients agree fully with attendants, about 85% of whom indicated that the stated behaviors are prohibited. In short, because the responses of mentally retarded individuals to yes-no questions are not entirely valid, the picture of the sample derived from interview data depends largely on how questions happen to be worded.

SLIGHT CHANGES IN QUESTION WORDING

Serendipity gave us some insight into the effects of question wording that we would not otherwise have had. A typographical error caused us to ask institutionalized adults two closely related questions, "Who decides what chores you do?" and "Who decides what chores to do?" The apparent effect of this change in question wording on responses was startling.

Responsiveness to these two open-ended questions was identical at 34.5%. Regarding consistency, the fact that the first of the two questions was repeated on the two forms of the interview gives us a baseline estimate of response reliability over a period of a week was 83.3%. For the two comparisons between the different wordings of the question, consistency figures were much lower - 62.6% and 73.3%. While the number of persons responding was small, the comparison of these fig-

ures with the 83.3% figure for test-retest reliability suggests that very small changes in question wording may have effects on clients' responses.

Agreement figures for the two "you do" questions were 88.2% and 75.0%. Agreement for the "to do" wording was considerably lower (70.6%). Possibly the "to do" question was more ambiguous.

Conclusions

Clearly, even very slight changes in question wording can have important effects on the reliability and validity of the responses of mentally retarded respondents.

USE OF QUANTITATIVE QUALIFIERS

Included in follow-up interviews with the severely retarded sample were five yes-no questions and one either-or question with picture alternatives, all dealing with whether or not the client makes his or her bed.

The questions, which used quantitative modifiers to introduce slight differences in meaning, were as follows:

Do you always make your bed here in the cottage?
Most days, do you make your bed here in the cottage?
Do you usually make your bed here in the cottage?
Do you sometimes make your bed here in the cottage?
Do you ever make your bed here in the cottage?
This picture shows a boy/girl who makes his/her bed here in the cottage. This picture shows a boy/girl who does not ever make his/her bed here. Which boy/girl is most like you?
point to the picture.

Responsiveness for the five yes-no questions ranged from 79.3% to

Table 8.1 % "Yes" Responses to Bed-making Questions

Question	% Yes	N
Always	96.0	25
Most days	100.0	26
Usually	96.2	26
Sometimes	92.0	25
Ever (Yes-no)	100.0	25
Ever (Picture either-or)	77.8	27

89.7%, averaging 84.4%. There is no indication that responsiveness differed systematically as a function of the quantitative modifier used. Responsiveness for the either-or question using pictures was 89.7%.

Regarding consistency, the best strategy for extracting meaningful information from these comparisons is a kind of hierarchical analysis; if a client responds "yes" to the "always" question, then he or she must necessarily respond "yes" to the other questions to remain consistent. Table 8.1 shows the percentage of "yes" responses to the six bed-making questions. As can be seen, no more respondents answered "yes" to the "usually" or "sometimes" question than to the "always" question. This calls into question the validity of responses. Although it is logically possible that clients virtually all make their beds every day, responses to the pictorial either-or question suggest otherwise. Because responses to these questions represented subtle quantitative differences, agreement with significant others is not a useful index of validity.

Conclusions

Perhaps the conclusion best justified by the data is that clients were relatively insensitive to the quantitative shadings involved in these questions - they tended to respond in the affirmative regardless of the quantitative modifier used. This suggests that laborious efforts to refine questions through quantitative modifiers may not have the effects that question designers intended them to have.

SPECIFIC VERSUS VAGUE WORDING

Two comparisons of open-ended follow-up questions in the institutionalized adult sample compared general or vague question wording with a more specific wording which restated the main content of the question. These comparisons suggested that type of wording had little effect on consistency or agreement, but had a clear effect on responsiveness; the specific wording maximized responsiveness. For example, as a follow-up to a question about attitudes toward living in the institution, clients were asked either "Why do you feel that way?" or "Why do you feel that way about living here?" responsiveness to the follow-up using the former wording was 26.2%, as compared to 54.1% for

the latter (more Specific) wording. Results were identical for a follow-up question dealing with clients' attitudes towards the food, with the responsiveness figure favoring the more specifically worded question, 43.9% to 30.6%. In short, it is clear that if maximizing responsiveness to open-ended follow-ups is the desired goal, the interviewer should use wordings which repeat the main content of the question. This strategy does not appear to lessen validity in any way, and it has the decided advantage of reminding respondents of the question content.

THE USE OF PROBES WITH OPEN-ENDED QUESTIONS

As the data presented thus far suggest, the researcher or program planner who wishes to interview mentally retarded individuals faces a dilemma: highly structured question formats such as yes-no questions maximize responsiveness but tend to introduce response biases which seriously reduce response validity; less structured formats such as open-ended questions, on the other hand, can be answered by few and elicit little information from them. In order to find ways to improve open-ended questions, we at times probed "what else" as respondents answered them, hoping to increase the amount of information gained. The following section is a comparison of responses to open-ended questions using such probes and to otherwise identical questions with no probing. Because the probe follows the first response, the formats are identical with respect to the percentage of clients responding appropriately to the question, and analysis of responsiveness is omitted from the following discussion.

Illustrative Comparison

Institutionalized adults were asked the questions "What do you usually do for fun when you are by yourself?" and "What do you usually do for fun when you are with someone?" (probe: "What else?" until client can think of no more).

Consistency of responding to these questions was evaluated in terms of both the number of responses given to each question and the content of the particular responses given. In terms of numbers of responses, clients averaged 5.0 responses to the questions with probe and 2.9 responses to the question without probe. Clearly, the use of the probe increases the number of responses given by the clients who could answer the questions.

Consistency of response content, averaged over the ten categories into which responses were coded, was 80.0%. However, most of this observed consistency occurred because clients did not mention a category in response to either format; only 8.6% of comparisons were consistent in that the same category was mentioned both times. Of more interest is the fact that most inconsistencies occurred when the client mentioned a category in response to the question with probe and failed to mention the same category when no probe was given. Of all comparisons, 15.0% (or three-fourths of all inconsistencies) were inconsistencies of this type.

Client-attendant agreement can also be evaluated either by the number of responses given, or by the response categories used. For the question without probe, clients averaged 2.5 responses versus 2.7 for attendants. For the question with probe, clients averaged 4.7 responses versus 2.6 for attendants. Thus, the use of the probe induced

clients to mention far more activities than did their attendants; this finding at least suggests that the probe technique may, by placing demands on the client, cause overreporting.

Agreement between attendants' and clients' response to the question without probe, averaged across the ten response categories, was 82.3%. For the question with probe, average agreement across the ten response categories was 74.6%. Thus, the use of the probe apparently reduced agreement, most disagreement occurring because the client mentioned a category which was not mentioned by the attendant. All client-attendant disagreements in the probe format (17.7% of all paired responses) were inconsistencies of this type. In short, the use of the probe reduced agreement by inducing clients to mention categories not mentioned by attendants.

Additional Comparisons

Institutionalized adults were asked three additional pairs of open-ended questions comparing the probing approach with identical questions without probing. These questions asked, "What do you and your friends usually do together?" "Where do you usually get your money?" and "What are they teaching you there?" a question which was asked only of respondents who had answered "Yes" when asked "Do you go to school or take classes?" Results from additional comparisons of the probe and no probe formats were remarkably similar. In each and every case, mentally retarded individuals gave more responses when the interviewer probed "What else?" after each response and in each case the use of the probe decreased agreement.

Conclusions

A remarkably clear pattern emerged from these comparisons. First, the use of the probe technique was effective in eliciting more responses from mentally retarded respondents. Unfortunately, this had the effect of reducing agreement. It appears that the probe created a demand which was met by clients, regardless of whether the information provided was valid or not. The net effect is that probing seems to cause inappropriate overresponding, reducing response validity. There was also some evidence that underreporting occurs when no probe is used. Possibly, the optimal approach would be to probe following one or two responses, rather than continuing to probe indefinitely.

The disastrous effects of probing ad infinitum were clear in several of the interviews. One severely retarded male, asked about what he and his friends usually do together, apparently felt compelled to repeat himself in response to each, "What else?" His responses were "play round," "play around," "go outside and play," and "go play," in that order. Clearly no new information was obtained; the only effect of probing was to force him to change his wording a little in hopes of convincing the interviewer that he had new news. Another, in response to the question, "Where do you get your money?" said, "I work." However, he then interpreted the "where else?" probes that followed as asking where else he worked, and then responded, in turn, "at the kitchen," "cottages," and "in here." Still others, interpreting the probes as a license to play word games, ran off lists of over 10 activities, clearly doing so to meet the interviewer's demands with each new probe. One, though, he repeated himself somewhat,

came up with the staggering total of 27 responses to 27 "What else's." His last response had to do with cleaning, and from then on he took the "What else's" to refer to other things he cleaned, giving the interviewer a much-welcomed opportunity to go on to the next question.

Multiple-Choice Questions

THE MEANING OF QUANTITATIVE RESPONSE OPTIONS

We used many multiple choice questions offering response alternatives like "a lot" and "some," which differ along a quantitative dimension, in order to collect information about extents of involvement in various activities. The many problems we encountered in the use of quantitative multiple choice questions in interviewing mentally retarded persons may be a result of the difficulty those individuals have in attributing consistent meaning to quantitative words and phrases. One comparison will serve to illustrate this phenomenon. We asked institutionalized adults "How many friends do you have? A lot, some, not many, none?" and simply "How many friends do you have?" Respondents were more responsive to the multiple choice question (55.2%) than to the open-ended question (39.7%).

Regarding consistency, the numbers offered in response to the open-ended item did not correspond well to the categories endorsed in responding to the multiple choice question. For example, the 64.7% of respondents who answered from one to six friends on the open-ended question answered anywhere from "none" to "a lot" when asked the multiple choice form. Among respondents who said they had seven or more friends in response to the open-ended question, responses to

the multiple choice form were divided between "a lot" and "not many."

In view of the fact that these quantitative terms appeared to have very diverse meanings among institutionalized adults, it is not surprising that clients and attendants were not in consistently close agreement. For example, in response to the multiple choice question, 18.8% of residents said they had no friends while their attendant said they had five or more friends; another 21.9% said they had "a lot" of friends while the attendant was reporting four or fewer friends. For the open-ended question, 17.4% of comparisons involved clients saying they had one to three friends, while the attendant reported over ten friends. Another 17.4% of comparisons involved clients saying they had no friends while the attendant said that they had at least one friend.

Conclusions

In short, the evidence suggests that the use of quantitative concepts by our respondents was very inconsistent, that the meanings of these terms apparently differed greatly from person to person. In view of this fact, the disappointing results achieved using quantitative multiple choice questions are not surprising. Sample 8.2 illustrates some of the difficulties our interviewees had with quantitative concepts.

ALTERNATIVE MULTIPLE CHOICE FORMATS

Quantitative multiple choice questions were included in our interviews largely in an attempt to get an idea of the extent of involvement in various activities on the part of our respondents. However, we were from the beginning concerned that four response alternatives might tend to confuse mentally retarded individuals. In

Sample 8.2

A Sampling of Problems with Quantitative Concepts

1. Q: How often do you leave here to go see your family: A lot, sometimes, not much, or never?

A: (Respondent said yes as each alternative was given.)

2. Q: How many people sleep in your bedroom?

A: Charlotte

3. Q: How many people sleep in your bedroom?

A: We got a whole bunch of beds back there. I can't count 'em all. Not right now.

4. Q: Are you usually happy or sad?

A: I'm always happy sometimes.

5. Q: Are you usually sad?

A: Yes, when somebody hits me I am. (On repetition of the question, he said, "No, not very much.")

6. Q: How much do you watch TV - a lot, some, not much or never?

A: Some, much.

Q: (repeated)

A: Never.

7. Q: Do you watch TV a lot, some or not much?

A: Some, not a lot, much.

A: (repeated)

A: I'm going to watch Elvis Presley tonight.

8. Q: How much do you read books, magazines, or newspapers -
Never, not much, some, or a lot?

A: I can't read. I'm sorry, but I can't.

Q: (Repeated since none of the four options was chosen)

A: A lots.

And finally, a reconsideration of knee-jerk acquiescence:

9. Q: Could you buy a new TV with \$10.00?

A: Yeah, I do. No, I can't. You can't buy a TV with that
money. You pulling my leg off.

Table 8.2: Yes-No Plus Three-Choice Multiple Choice vs. Four-Choice Multiple Choice

Content	Four-Choice Multiple Choice		Yes-No & Three Choice Multiple Choice		Consistency
	% RESP	% AGREE	% RESP	% AGREE	% Consistent
How often do you leave here to go see your family?	78.0%	50.0%	66.7%	52.4%	77.3%
How often does anybody in your family come to see you?	75.6%	30.0%	69.0%	22.4%	48.0%
How much do you watch TV?	55.2%	22.2%	56.9%	27.3%	44.4%
How often do you go to church?	48.3%	70.4%	61.4%	70.6%	85.2%
How often do you go to stores?	49.1%	NA	60.3%	15.0%	48.0%
How much do you listen to the radio or record player?	41.3%	45.8%	55.2%	42.5%	52.2%
How often do you go out to the movies?	49.1%	5.9%	50.0%	NA	29.2%
How much do you read books, magazines or newspapers?	52.6%	18.1%	48.3%	32.0%	50.0%
How often do you go out to eat?	55.2%	16.1%			33.3%
Average:	56.0%	32.3%	57.6%	37.5%	52.0%

an effort to simplify the task demands but still obtain quantitative information, we combined some yes-no questions with three-choice multiple choice questions asked of only those respondents who answered the yes-no question "yes." For example, we asked "Do you leave here to go see your family? (if yes) Do you go see them a lot, sometimes, or not much? The comparisons that follow compare the four-choice multiple choice format with the yes-no followed by contingent three-choice multiple choice format. In the analysis that follows, responsiveness figures have been transformed so that responsiveness for the two formats can be directly compared, and consistency and agreement figures have been similarly adjusted (with a "no" response to the yes-no question coded as a "never" response to the four-choice multiple choice question).

Illustrative Comparison

One comparison of these two formats consisted of asking institutionalized adults both "How much do you read books, magazines, or newspapers? Never, not much, some, or a lot?" and "Do you read books, magazines, or newspapers? (if yes) Do you read books, magazines, or newspapers a lot, some or not much?"

A total of 48.3% of respondents responded appropriately to the yes-no question and subsequent three-choice multiple choice question. This is slightly less than the 52.6% who gave appropriate answers to the four-choice question.

Of 26 subjects who responded appropriately to both formats, only 50.0% gave consistent answers. The pattern of inconsistency is of interest: in response to both questions, the response alternative most frequently chosen was the last-

mentioned. Thus, 11.5% of comparisons were inconsistencies due to the client choosing the last option in both cases. This suggests that consistency was reduced in part by a last option response bias. In addition, 80.8% of clients responded "yes" to the yes-no question, indicating that they at least sometimes read. In response to the four-choice multiple choice question, only 69.3% of clients indicated that they at least sometimes read. This contrast reflects acquiescence in response to the yes-no format.

The four-choice multiple choice question produced only 18.1% agreement with attendants. Most disagreements (50.5% of comparisons) occurred because the client responded "a lot" (the last option) and the attendant said something less. Agreement for the yes-no plus three-choice multiple choice question was slightly higher at 32.0%. Of all comparisons with attendants, 24.0% were "not much" responses from the client (the last option), coupled with a response of either "never" or "some" by the attendant. In short, agreement is low for both formats, with the pattern of inconsistencies indicative of some last option response bias on the part of the residents.

Additional Comparison

A total of nine comparisons of four-choice multiple choice questions with yes-no plus three-choice multiple choice questions were conducted. Table 8.2 depicts responsiveness, consistency, and agreement data for all nine comparisons. Although the specific comparison discussed above is generally representative of these sets of questions, it is clear that there is a lot of variability within the set. Nevertheless, the general conclusions suggested by the "read" ques-

tions are unaltered; specifically, neither format appears attractive. Responsiveness, consistency, and agreement are poor for both formats, neither appears to offer a definite advantage over the other. Agreement in particular was poor for both formats 37.5% for the yes-no plus follow-up approach versus 32.3% for the four-choice questions. Only the questions about church attendance appeared to work well, possibly because there is a relatively clear external standard suggesting that once a week or more is "a lot."

It should be pointed out that the last option response bias discussed above was observed frequently but not inevitably within this set. It appears that mentally retarded respondents are often biased in favor of the last-mentioned alternative, but that this set is weak enough that it is frequently overshadowed by other factors. The important point is that these quantitative multiple choice questions do not yield useful information.

VERBAL MULTIPLE CHOICE QUESTIONS VS. MULTIPLE CHOICE QUESTIONS WITH PICTURES

A direct comparison of standard multiple choice questions with multiple choice questions offering pictures as response alternatives was undertaken to see if one format had a definite advantage. The use of pictures as response alternatives was primarily intended to elicit information from clients with few verbal skills. These comparisons involved subjective questions.

Illustrative Comparison

Institutionalized adults were asked, "How do you like living here: a lot, some, not much, or not at all?" and "Which picture shows how

you feel about living here? Point to the picture." (Sad face followed by somewhat sad, somewhat happy, and happy faces.)

Consistent with prior observations regarding responsiveness to picture formats, the use of pictures increased responsiveness. In response to the picture alternatives, 70.7% of clients responded appropriately by pointing; for the verbal multiple choice question, responsiveness was 50.9%.

Responding across the two formats of the multiple choice question was very inconsistent; the overall consistency figure was only 21.4%. As to the pattern of this inconsistency, it appears that the use of pictures tended to spread responses across response alternatives. Responding to the verbal format was biased toward the "a lot" end of the continuum, with 35.7% of clients replying "a lot" (the first option). Responding to the picture format did not favor one end of the response continuum, as 32.1% of residents answered "a lot" and 39.3% replied "not at all." This could reflect a first option response bias in the case of verbal alternatives, or simply a greater tendency toward random responding in the case of picture alternatives. Since a bias toward choosing the last option on multiple choice questions has been observed elsewhere, we favor the latter interpretation.

Although these questions were subjective, we asked attendants to characterize residents' satisfaction with the institution. Agreement for the picture format was 21.3%. Agreement for the word format was slightly higher, at 32.3%. In both cases, inconsistencies arose primarily because attendants rated clients as having more positive feelings than were reported by

the clients themselves. This may well indicate that there is a real difference of opinion between clients and attendants on the issue of satisfaction with the institution. However, consistency data still suggest that neither of these formats successfully elicited valid answers, and that the picture format, while it enabled many persons to respond, is particularly suspect.

Additional Comparisons

Two additional comparisons of picture and verbal multiple choice questions were conducted, but without data on agreement with attendants. These questions involved satisfaction with the food and with the people at the institution. In both additional comparisons, the use of pictures was associated with higher responsiveness than was the verbal questions. These comparisons were 67.9% vs. 50.0%, and 67.9% vs. 55.2%, respectively. Consistency figures for the questions regarding food and people were 33.3% and 37.8%, respectively. This very low consistency, indicative of a lack of reliability of responses to these quantitative multiple choice questions, suggests that neither format successfully generated reliable information.

Conclusions

Clearly, the use of pictures as response alternatives increased responsiveness to these multiple choice questions. However, consistency was so low as to suggest that neither format is very useful. Based on agreement data and the pattern of inconsistencies observed, it appears that the gain in responsiveness achieved by using pictures is not accompanied by an increase in reliability and validity. Sample 8.3 illustrates some of the problems with these picture choice questions.

These discouraging data are not surprising in view of the fact (discussed in Chapter 6) that test-retest reliability for one group of picture multiple choice questions averaged only 46.2%.

can be compared roughly with 70.0% reliability for verbal multiple choice questions on different topics, as reported in Chapter 6.) Responses of mentally retarded persons to quantitative multiple choice questions, whether verbal or pictorial, appear to be inherently unreliable.

DISCRETE MULTIPLE CHOICE QUESTIONS

The multiple choice questions previously discussed, which appear unsatisfactory for use with mentally retarded populations, involved subjective judgments on the part of respondents and used response alternatives which differed along a quantitative dimension. Another group of multiple choice questions in our interviews dealt with factual information and offered response alternatives which differed in a qualitative manner--that is, discrete response options. Discrete multiple choice questions in both "point to the picture" and verbal response format compared directly with open-ended questions soliciting the same information.

Illustrative Comparison

Community children were asked how they traveled to school using each of the following three approaches: "Here are some ways people get to school. (Pointing to pictures). They take the bus, somebody drives them to school in a car, they walk, or they ride to school on their bicycles. Which way do you get to school most days? Point to the picture."; "Most days, how do you get to school?";

Sample 8.3

and Results with Picture Choice Questions

Failure Story:

Q: Which picture shows how you feel about living here?

Point to the picture.

A: (pointed to "not at all" saddest face of four faces)

Q: Why do you feel that way?

A: I don't like to feel that way.

Q: (repeated)

A: I don't feel that way.

Weird Story

Q: Which picture shows how you feel about the people here? Point to the picture.

A: (did not point) They're not good. They. . . (stopped)

Q: (repeated)

A: (did not point) Cause they get mad about the way she acts - she'll do anything I say (reference not clear)

Q: Why do you feel that way?

A: They don't assume I harp on me to say. They don't understand what I say.

Success Story:

Q: Which picture shows how you feel about the people here.

Point to the picture.

A (pointed to "not at all" saddest face)

Q: Why do you feel that way?

A: He looks like he's crying, tears falling from his eyes.

People were making fun of him. He wants to go home to his parents.

Note: Unfortunately this success story is the exception.

Typically respondents had difficulty giving consistent responses to these picture choice questions about satisfaction.

and "Most days, do you get to school on the bus, in a car, by walking, or riding a bicycle?"

Responsiveness to the picture multiple choice question was 100%; all children asked the question were able to respond appropriately by pointing to one of the pictures. For the verbal multiple choice question, responsiveness was 94.5%. The open-ended question elicited the lowest percentage of responses from children (73.6%). Clearly, the use of multiple choice questions in general and of picture multiple choice in particular optimized responsiveness.

Consistency for comparisons using the discrete multiple choice format was exceptionally high. Comparison of the picture multiple choice format with the open-ended question produced a consistency figure of 95.2% for 41 response pairs. In the comparisons of the verbal multiple choice and open-ended questions, 95.1% of children responded consistently. The comparison between the two multiple choice questions produced a consistency figure of 86.9%. For the latter two comparisons, most inconsistencies arose because children often chose the last option in the verbal multiple choice format, again suggesting a last option bias. However, this interpretation must be tempered by the realization that consistency was so high that patterns of inconsistency involved very few subjects; the major conclusion remains that consistency was remarkably high for these questions.

Parent-child agreement figures for the picture multiple choice (85.5%) and the open-ended question (85.4%) were almost the same. There was slightly less agreement between parents and children on the verbal

multiple choice format (77.4%). The only pattern observed in the disagreements was a slight last option bias on the verbal multiple choice question; 13.2% of pairs disagreed because the child said "bike" (the last option) while the parents said either "car" or "bus." In this case, then, the picture multiple choice question is preferable on grounds of both responsiveness and validity.

Additional Comparisons

There was another comparison, also involving community children, of four questions asking in what kind of place the child lived (house, apartment building, trailer house, or duplex). The questions were discrete multiple choice, picture and verbal, with response alternatives presented in original and alternative orders as a test for order effects.

Average responsiveness for the picture choice questions was 96.5%, as opposed to 81.8% for the verbal questions. Consistency of response between the two verbal choice formats was slightly higher than that for the picture questions--85.7% as compared to 78.2%. For the two comparisons between word and picture formats presenting options in the same order, consistency averaged 81.8%. The two comparisons between verbal and picture formats which presented options in different orders yielded a similar average consistency figure of 78.8%. These consistency data suggest that (1) the superiority of the picture questions observed earlier was not supported, and (2) order biases are not particularly important in these formats, since different order comparisons are about as consistent as same order comparisons. Average agreement between parents and their children for the verbal choice questions was

81.8%, while average agreement for the picture format questions was 70.5%. This again reverses the conclusion drawn from the transportation questions, but it is not clear why the validity of responses to the picture choice question was lower here.

Conclusions

One conclusion which can be drawn from these comparisons is that discrete multiple choice questions about basic facts differ strikingly from their quantitative counterparts. These questions yielded high responsiveness, high consistency, and high agreement relative to most other formats. Comparisons between picture choice and verbal choice approaches resulted in inconsistent findings, so that a global judgment of their relative merits is impossible. There is some evidence for a last option bias in verbal multiple choice questions, but this phenomenon is unstable. To the extent that order bias affects verbal choice questions, it at least does not consistently reduce their validity below that of picture choice questions. However, since picture multiple choice questions consistently maximize responsiveness and there is no reason to believe that they are consistently less valid than verbal choice questions, picture multiple choice would appear to be an excellent method for eliciting factual information from mentally retarded respondents. Unfortunately, we did not test the discrete multiple choice format with other content areas, so it is impossible to determine whether the success found here would generalize to other topic areas, or whether retarded children simply have solid enough understanding of how they get to school and where they live that almost any format would work.

Yes-No Vs. Open-Ended Questions

LISTING QUESTIONS

One type of information which we sought to obtain from mentally retarded persons was enumerations of activities, problems, interests, and so on. The decision faced by an interviewer interested in obtaining such information hinges on the issue of question structure. Should one choose highly structured formats and run the risk of response biases such as acquiescence or should one choose open-ended formats and run the risk that few interviewees will be able to answer the questions? To evaluate this problem, we directly compared open-ended questions asking for active enumeration of responses with sets of yes-no questions intended to obtain the same information.

Illustrative Comparison

One comparison of these two formats, drawn from the institutionalized children sample, accurately represents finding for the group of comparisons as a whole. Institutionalized children were asked eight yes-no questions about sports in which they might participate (for example, "Do you play football?"). Residents were also asked an open-ended question soliciting the same information: "Do you play any sports?" (if yes) Which ones? Responses to the open-ended format were subsequently coded into categories corresponding to the yes-no questions, so that each client responded "yes" or "no" when asked if he or she plays a given sport and was categorized as having mentioned or not mentioned that sport in response to the open-ended question. Attendants were asked a parallel list of yes-no questions ("Does this client play

any of these sports: Football? Baseball? and so on).

Average responsiveness to the eight yes-no questions was 85.6%. This compares favorably with the 55.2% figure for the open-ended question. Consistency was evaluated by averaging consistency across the eight pairings of a yes-no question with its corresponding response category. Average consistency was 49.7%, with inconsistency resulting almost exclusively from pairings in which the client replied "yes" on the yes-no format but failed to mention that response category in response to the open-ended question. Of all comparisons, fully 48.1% were inconsistencies of this type. As well as illustrating the ubiquitous impact of acquiescence, this finding highlights the possibility that underreporting is a serious problem in the response of mentally retarded persons to open-ended questions.

Examination of data regarding agreement with attendants is necessary to determine which effect is stronger.

Client-attendant agreement was calculated for the eight-yes-no questions and for the eight categories of response to the open-ended questions. Average agreement for the yes-no questions was 52.1%, as opposed to 60.1% for the open-ended question. For the yes-no questions, 42.3% of all pairings were disagreements in which the child said "yes" while the attendant said "no," and an additional 5.6% were disagreements in which the child said "no" and the attendant "yes." Again, the overwhelming impact of acquiescence is evident. For the open-ended question, 31.9% of client-attendant pairs disagreed in that the attendant said "yes" regarding a sport not mentioned by the retarded respondent. Only 8.1% of pairings were disagreements in

which the child mentioned a sport not indicated by the attendant. Thus, retarded interviewees tended to underreport their activities in response to the open-ended question. In short, acquiescence leads to overreporting in response to yes-no lists and clients are prone to underreport in response to open-ended questions; the latter bias may be slightly less dramatic.

Additional Comparisons

Four additional sets of questions, asked of institutionalized adults and institutionalized children, compared open-ended and yes-no listings questions. These sets of questions solicited information about the client's problems, actions of others that bother the client, nice things people do for the client, and participation in arts and crafts.

In each comparison, far more clients gave appropriate responses to the yes-no questions than the open-ended questions. The responsiveness figures for the four yes-no questions ranged approximately from 60-80%, while those for the open-ended questions ranged from 20-50%. The superiority of yes-no questions in eliciting appropriate responses is clear.

As with responsiveness, the consistency data for these comparisons are remarkably uniform and echo the results from the illustrative comparison. That is, the two formats generate very different pictures of our clients, because clients claim far more activities, problems, and so on in response to the yes-no format than when asked open-ended questions. Consistency for these comparisons varied, but in all cases the largest source of inconsistencies consisted of "yes" responses to the yes-no format coupled with

"not mentioned" responses to the open-ended question.

Only one of these comparisons, contrasting an open-ended question asking the client to enumerate problems that he or she has with a series of yes-no questions about particular problems, included attendants' responses. Agreement was far lower for the yes-no than the open-ended questions (average agreement of 45.6% and 92.4%, respectively). The yes-no format again led to overreporting because of acquiescence, as 44.9% of comparisons were disagreements in which the client answered "yes" and the attendant answered "no." The fact that clients clearly acquiesced even in response to these questions, for which "yes" responses were not socially desirable responses, indicates that in at least some cases acquiescence per se can overrule the influence of social desirability. The open-ended format again appeared to lead to underreporting, as all disagreements (7.5% of pairs) occurred when an attendant said "yes" regarding a problem not mentioned by the client. The very high agreement on this question resulted from the fact that neither clients nor attendants mentioned many problems; in 89.1% of pairings, the two agreed in that neither indicated a given problem.

Conclusions

These comparisons are an excellent illustration of the difficulties of interviewing mentally retarded individuals, as the use of two different methodologies to seek the same information yielded completely dissimilar pictures of the target population. Yes-no questions have the advantage of high responsiveness, but are clearly subject to strong systematic bias (acquiescence). Low agreement for

these questions suggests that acquiescence leads mentally retarded persons to overreport whatever they are being asked about. Open-ended questions have the advantage of relatively high agreement, but generate very low responsiveness so that information drawn from open-ended questions represents only a selected fraction of the population. Further, the number of items enumerated in response to these questions is low even among those who respond appropriately; the very low proportion of mention-mention agreements in these comparisons raises the possibility of underreporting by both mentally retarded respondents and their attendants. The truth probably lies somewhere between these extremes of underreporting and overreporting. Unfortunately, it seems impossible to establish exactly where, though our findings hint that overreporting in response to yes-no questions may be the more pronounced bias.

DECISION-MAKING QUESTIONS

Illustrative Comparison

Two open-ended questions, "Who decides what chores you do?" and "Who decides what chores to do?" were compared with a yes-no question, "Do you decide what chores to do?" The first open-ended question was asked on both forms, allowing for three comparisons of the open-ended and yes-no formats. These questions differ fundamentally from open-ended questions previously discussed, in that they do not solicit enumerations. Rather, the client's response was coded into one of two categories (client decides vs. someone else decides). Thus, the information demanded by these questions is closely comparable to that solicited by either-or questions.

Responsiveness for the three open-ended questions ranged from 32.8% to 34.5%, averaging 33.9%. This figure was clearly exceeded by the 54.4% responsiveness for the yes-no question; the superiority of the yes-no format in eliciting appropriate responses emerges with monotonous regularity.

Consistency for the three comparisons of the yes-no questions with open-ended questions ranged from 56.3% to 68.9%, averaging 63.9%. Examination of the data reveals a definite pattern in the inconsistencies. Of all comparisons, 32.4% were errors in which the client responded "someone else" to the open-ended question, but answered "yes" (client decides) when asked the yes-no form of the question. This fraction accounts for nearly all the inconsistencies; clearly, acquiescence powerfully influenced the responses of clients to the yes-no format. As a result, one would conclude on the basis of the yes-no question that 52.1% of the clients decide for themselves, whereas the three open-ended questions yield a much lower figure (23.4%).

Agreement with attendants differed radically for the yes-no and open-ended questions; agreement for the yes-no question was 34.6%, while agreement averaged 77.9% for the three open-ended questions. For the yes-no question, 65.4% of comparisons were disagreements in which the client responded "yes" (client decides), while the attendant responded that someone else decided. Again, acquiescence operated to reduce the validity of responses of retarded adults to the yes-no question.

Additional Comparison

A similar comparison of yes-no and open-ended questions, also conducted with institutionalized

adults, involved decisions about spending money. The results of this comparison are virtually identical to those of the previously discussed comparison.

Conclusions

As always, responsiveness for yes-no questions was far higher than for open-ended questions. However, in these comparisons the reliability and validity of the open-ended questions were superior. Client-attendant agreement was in each case much higher for the open-ended questions, as acquiescence apparently reduced the validity of responses to the yes-no questions. It appears that open-ended questions soliciting information which can be coded into a few present categories may be quite useful, as long as one is willing to make a sacrifice in responsiveness.

Yes-No Vs. Either-Or Questions

On another test of alternative formats, we compared sets of yes-no and either-or questions dealing with the same topics. A clear comparison of these formats requires an understanding of the strengths and weaknesses of each questioning strategy taken separately. Thus, we begin by analyzing responses to sets of oppositely worded yes-no questions and to sets of either-or questions with response alternatives presented in reverse order.

OPPOSITELY WORDED YES-NO QUESTIONS

Two comparisons of oppositely worded yes-no questions will be discussed below. Responsiveness is unaffected by reversing the wording of yes-no questions and is not analyzed here.

Table 8-3: Oppositely Worded Yes-No Questions

Does somebody else decide how you spend your money?	Do you decide how to spend your money?		
	No	Yes	Total
No	11.8%	26.5%	39.3%
Yes	11.8%	50.0%	61.8%
Total	23.6%	76.5%	

Consistency = 38.3%

Illustrative Comparison

Institutionalized adults were asked "Do you decide how to spend your money?" and "Does somebody else decide how to spend your money?" The contingency table for comparison of the two questions is depicted in Table 8.3. Only 38.3% of the adults, far below the 50% which would be expected if responding to both questions were random, gave consistent responses. While 76.5% said "yes" to "Do you decide...?" 61.8% also said "yes" to "Does somebody else decide...?" Fully 50.0% of the response comparisons were self-contradictions in which the client responded "yes" to both forms of the question. Acquiescence is clearly a major determinant of responses to these questions.

For the "Do you decide...?" question, agreement with attendants was 55.5%. For the oppositely worded question, agreement was 54.3%. In both cases, the majority of disagreements consisted of a "yes" response from the client, paired with a "no" response by the attendant, again indicating the importance of acquiescence by mentally retarded persons. Another set of oppositely-

worded yes-no questions, asked of community children, produced virtually identical findings.

Conclusions

Both of these pairs of oppositely worded yes-no questions clearly illustrate the enormous importance of acquiescence among mentally retarded persons. These respondents tend to answer "yes," even when such responses are self-contradictory. Additionally, clients tended to respond "yes" far more often than significant others. Clearly, acquiescence effects are powerful enough that the validity of retarded persons' responses to yes-no questions is minimal.

REVERSED EITHER-OR QUESTIONS

Illustrative Comparison

A number of pairs of either-or questions were asked, such that the members of the pair dealt with identical content but the order of presentation of response alternatives was reversed. One such comparison in the children's community sample will be discussed below. The questions were "Most days, does somebody else tell you what to

wear to school, or do you decide?" and "Most days, is it up to you to decide what to wear to school or does somebody else tell you what to wear?"

Responsiveness, 63.6% to the first either-or question and 65.5% to the second, was unaffected by the change in wording. Consistency was impressive for this pair of questions, in that 83.3% of children who responded appropriately to both questions responded in a consistent manner. The type of inconsistency was evenly divided, suggesting a lack of position preference in response to this question.

Agreement with parents was fairly high for both forms of the question. Agreement was 72.5% for the question in which "somebody else" was first alternative, and was 73.2% for the other question. For the former question, types of disagreement were evenly mixed; for the latter, most inconsistencies (19.5% of comparisons) involved the choice of the last option by the child. This constitutes evidence for a weak position bias in favor of the last-mentioned alternative.

Conclusions

This pair of either-or questions generated relatively high consistency and parent-child agreement. There was some evidence of a weak position bias by mentally retarded persons, but little indication of powerful response effects. To this point, then, either-or questions appear superior to yes-no questions, but head-to-head comparisons of the two formats are needed before we can be sure.

YES-NO VS. EITHER-OR: FOUR-QUESTION SETS FROM INSTITUTIONALIZED ADULTS

Two sets of questions asked of institutionalized adults, comparing either-or and yes-no questions, were repeated in two interviews a week apart, yielding a measure of test-retest reliability as well as consistency measures.

Illustrative Comparison

One of these four-question sets inquired about the client's happiness, using the following questions:

- Are you usually happy?
- Are you usually sad?
- Are you usually happy or sad?
- Are you usually sad or happy?

All questions were repeated on both administrations of the interview, so that a total of eight questions is involved in this comparison.

Average responsiveness for the eight questions was 65.5%. Responsiveness ranged from 58.6% to 70.7%, and was slightly higher for the four yes-no questions than for the either-or questions--68.1% vs. 62.8%.

One measure of consistency compared each question with its identical counterpart on the other form, giving a measure of test-retest reliability. For the "happy" and "sad" yes-no questions respectively, consistency figures were 97.3% and 60.5%. For the two either-or questions, consistency figures were 87.9% and 94.1%, resulting in average consistency for the four comparisons of 85.0%.

One point of interest is that consistency was very discrepant for the two yes-no questions. It seems likely that this is an effect of acquiescence; for the "happy" wording, clients' "true" response and their response set operated together, so that "yes" responses were firmly established. On the "sad" wording, response bias

and the "true" response operated in opposite directions. This conflict apparently introduced a degree of uncertainty which caused some clients to change their responses from interview to interview. Thus, the test-retest reliability of the yes-no questions appears to have depended in part on the wording of the questions. In contrast, consistency for the either-or questions differed only slightly as a function of wording. Test-retest reliability for both these types of questions was reasonably high. It was somewhat higher for either-or questions (on the average, 91.0% vs. 78.9%), probably because response set was of less importance for the either-or format.

Another type of consistency involved comparison of oppositely worded questions within a given format. For the four comparisons of oppositely worded yes-no questions asked on the two occasions, consistency averaged only 53.1%, ranging from 46.2% to 55.6%. Clearly, acquiescence was a major determinant of response for these questions, as on the average, 43.9% of clients contradicted themselves by responding "yes" to both questions. For comparisons of oppositely worded either-or questions, consistency ranged from 77.8% to 99.7%, averaging 86.2%. This relatively high, uniform consistency attests to the usefulness of either-or questions in interviewing mentally retarded individuals. However, there was some indication of a last option position bias inherent in responses to either-or questions. Clients tended overwhelmingly to respond "happy" to these questions, but this tendency was more pronounced when "happy" was the last mentioned response alternative (93.4% reported being "happy" when "happy" was the last choice vs. 81.8% when it was not). This response bias, however,

was much less powerful than was the acquiescence in response to yes-no questions.

This set of questions allowed for sixteen comparisons of yes-no to either-or questions. The eight comparisons of the question "Are you usually happy?" with either-or questions yielded an average consistency figure of 96.7%, ranging from 82.9% to 100.0%. These comparisons tended to be very consistent because a large majority of clients responded "yes" to the yes-no question and "happy" to the either-or format; an average of 87.4% of comparisons were consistencies of this type. However, for the eight comparisons of the question "Are you usually sad?" with either-or questions, consistency averaged only 57.3%. For this set, an average of 40.0% of format comparisons were inconsistencies in which the client responded "yes" to the yes-no question but answered "happy" when asked the either-or question. Clearly, "happy" was generally the "true" response of mentally retarded individuals to these questions, and responses to the either-or questions tended to reflect that fact (in spite of some order preference). However, when the yes-no question was worded so as to bring tendencies toward acquiescence and the "true" response into conflict, acquiescence caused many clients to inaccurately report that they are usually sad. The striking difference between the 87.4% and 57.3% average consistency figures for comparisons of either-or questions with the two yes-no questions attests to that fact.

Average agreement with attendants for the either-or questions was 85.8% and ranged from 79.3% to 94.3%. Average agreement for the yes-no questions was somewhat lower at 72.1%. This reduction in agreement is ac-

counted for by the very low agreement for the yes-no question "Are you usually sad?" which was 50.0% and 50.7% on the two forms. For this question, an average of 43.7% of client-attendant pairs were inconsistent in that the client said "yes" (sad) and the attendant indicated that the client was usually happy. Again, the major conclusion is that acquiescence largely invalidates the responses of mentally retarded respondents to yes-no questions.

Additional Comparison

Another comparison of yes-no and either-or questions, including measures of test-retest reliability, was conducted with institutionalized adults. This set of questions asked whether clients usually spend their time alone or with other people. Average responsiveness for the eight questions was 63.2%, and was slightly superior for the yes-no questions (68.3% vs. 58.0%). Test-retest reliability for the set averaged 76.2%. As in the previous set, the reliability of yes-no questions was augmented when acquiescence set and the clients' "true" answer coincided and was diminished when they were at odds.

Response consistency was again abysmally low when the oppositely worded yes-no questions were compared, averaging 52.0%. The most frequent response combination (43.4% of comparisons) had clients contradicting themselves by responding "yes" to both questions. Comparisons between oppositely worded either-or questions averaged 75.1%, suggesting superior validity for either-or questions. Again a last option response bias was observed, as the most frequent inconsistent response combination (18.3% of all comparisons) involved the client choosing the last option both times, as op-

posed to only 6.6% of clients who chose the first option both times.

Comparisons of yes-no and either-or questions did not reveal the radical differences in consistency figures for the two yes-no questions that characterized the "happy-sad" set. For the eight comparisons of either-or questions with "Are you usually with other people?" consistency averaged 69.5%. For the question "Are you usually by yourself?" consistency averaged 68.7%. This probably reflects the fact that answers were much more balanced between the two response alternatives for the alone-with others questions than for the happy-sad questions.

Average agreement for this set was 60.8% for the yes-no questions and 60.3% for the either-or questions. Again, agreement figures for the two yes-no questions were discrepant, averaging 51.4% for "Are you usually by yourself?" and averaging 70.2% for "Are you usually with other people?" For the either-or questions, agreement was slightly higher for the question for which the most frequently chosen response alternative was the last mentioned, 63.5% as opposed to 55.2%. Thus, client-attendant agreement is an overestimate of response validity when the "true" response and client response bias happen to coincide. This effect is apparent in the either-or comparisons, where agreement is greater when the "true" answer (with others) is the last-mentioned alternative. However, it is far more notable for the yes-no questions, where agreement is much higher when "yes" indicates that the client is usually with other people.

Conclusions

Both sets of questions suggest that either-or questions compare favorably with yes-no questions

as a method of gaining information from mentally retarded persons. Both either-or and yes-no questions involve some response biases which reduce the validity of responses. However, the acquiescence activated by yes-no questions is a crippling deficiency, whereas the position bias inherent in the either-or format is merely annoying. In the second set of questions discussed, the advantage of the either-or format is not so great. Nevertheless, consistency for the either-or questions in both cases approaches the test-retest reliability for these questions, and even in the second case, either-or questions are less affected by response set.

SUPPLEMENTARY YES-NO VS. EITHER-OR COMPARISONS

The set of yes-no versus either-or comparisons described above constitutes only a fraction of the data we collected regarding the relative merits of these two formats. Four separate four-question sets, similar to the above examples, were included in interviews with institutionalized children. In addition, two direct comparisons of isolated either-or and yes-no questions were conducted in interviews with institutionalized adults. Although some variation in patterns of response to these two formats was of course observed, the consistency with which these data characterized yes-no and either-or questions was remarkable. Specifically, in every case the responsiveness data slightly favored the yes-no format. However, in every comparison the evidence regarding validity unequivocally recommended the either-or questions, usually because of the powerful effect of acquiescence on the yes-no format. Again we must conclude that there is no point in eliciting invalid responses; responsiveness must yield to validity as a more essential cri-

terion. Direct comparisons of yes-no and either-or questions suggest that the either-or format is preferable as a means of gaining information from mentally retarded persons.

The Use of Pictures With Yes-No Questions

In a continuing effort to maximize the amount of information that we could gain from clients with low verbal skills, we asked some yes-no questions which included pictures to clarify the question. Two sets of questions directly compared yes-no questions using pictures with the same yes-no questions not using pictures. These comparisons can tell us whether the use of pictures in any way serves to reduce the acquiescence so commonly observed in response to yes-no questions.

Illustrative Comparison

The first such set, asked of institutionalized children, involved indoor games. The children were twice asked four yes-no questions such as "Do you play checkers?" once with words only and once including pictures of people engaged in the game named.

Responsiveness, 84.4% for the question with words only and 82.2% for the words plus pictures format, was apparently unaffected by question type. Respondent consistency for the four pairs of yes-no questions averaged 89.8%. Children tended overwhelmingly to respond "yes," with the result that most comparisons (75.1%) were consistencies in which the child responded "yes" both times.

The agreement data contain the most useful information regarding this format comparison. These data are summarized in Table 8.4. Overall agreement averaged only 42.4%. By far the greatest fraction of dis-

Table 8.4: Yes-No Questions with Words Plus Pictures vs. Words Only

Child-Attendant Agreement							
Question Type	Topic	N	Total% Agreement	Yes Agreement	No Agreement	Child-Yes S.O. - No	Child-No S.O. - Yes
No pictures	Checkers	45	53.3	42.2	11.1	42.2	4.4
	Cards	44	52.3	50.0	2.3	40.9	6.8
	Dominoes	43	32.6	23.3	9.3	55.8	11.6
	Other games	41	51.2	39.0	12.2	29.3	19.5
Group Average		43.3	47.4	38.6	8.7	42.1	10.6
Pictures	Checkers	43	62.8	44.2	18.6	34.9	2.3
	Cards	42	57.2	54.8	2.4	38.1	4.8
	Dominoes	43	32.6	23.3	9.3	55.8	11.6
	Other games	38	44.7	34.2	10.5	34.2	21.1
Group Average		41.5	48.3	39.1	10.2	40.8	10.0
Overall Average		42.4	48.3	38.9	9.5	41.4	10.3

agreements consisted of "yes" responses by clients coupled with "no" responses by attendants (41.1%, on the average), indicating powerful acquiescence effects. More important is the observation that the two formats differ little; they are characterized by equally low agreement and equally strong response biases.

Additional Comparison

A similar comparison of yes-no questions using pictures was conducted in the children's institutional sample, where children were asked if they engaged in ten different chores. Responsiveness averaged 69.8% for the ten pictures questions, vs. 70.9% for the words only format. Consistency of responding between formats, averaged across the ten questions, was 87.9%. This consistency tended overwhelmingly to result from the client answering "yes" to both formats; 74.1% of all comparisons were consistencies of this type. For the picture format, client-attendant agreement averaged 72.1%, resulting mostly from cases in which both client and attendant answered "yes." Most disagreements, 20.4% of all comparisons, occurred because the client responded "yes" and the attendant responded "no," indicating acquiescence. This pattern of responses was virtually repeated in the words only format, where overall agreement was 69.5%.

Conclusions

Yes-no questions using pictures tend to be invalidated by acquiescence, as are similar questions using words only. The use of pictures does not increase responsiveness or in any way improve the information obtained using yes-no questions. In short, the use of pictures offers no advantage over the simple words only format.

The Limits of Acquiescence

By now, the reader may be weary of our continual references to the problem of acquiescence. Yet it is of such importance that we feel a need to draw together some of our findings and present still more so that the implications of acquiescence are clearly appreciated.

One of the pairs of oppositely worded questions that were used in all three major samples was "Are you usually happy?" versus "Are you usually sad?" Of those persons who could respond to both questions, fully 50.9% of the community children, 43.8% of the institutionalized children, and 41.7% of the institution adults acquiesced, that is, answered "yes" to both questions. The correlations between acquiescence to this pair of questions and IQ in these samples were $-.30$, $-.40$, and $-.41$, respectively, all significant at the .05 level and all suggesting that lower IQ respondents are more likely than higher IQ respondents to acquiesce. The oppositely worded questions "Are you usually by yourself?" and "Are you usually with other people?" were also used with all three samples. Here, 50.0% of the community children, 39.4% of the institution children, and 41.7% of the institution adults contradicted themselves by saying "yes" to both questions. The correlations between acquiescence and IQ were $-.24$, $.03$, and $-.04$, respectively, with only the first statistically significant. Thus, in all three samples and on both pairs of questions, acquiescence rates were alarmingly high. And although lower IQ respondents tended to acquiesce more, the correlations were not large enough or stable enough to suggest that severely retarded persons are the only culprits. Instead, acquiescence was a pronounced problem of all levels of retardation.

Still, having found some relationship between acquiescence and IQ, we explored the matter further in new interviews with 29 moderately and severely retarded institutionalized children and adults who had proven to be largely, but not totally, unresponsive in their original interviews. We included the "happy-sad"

and "alone-with other" questions as well as other oppositely worded pairs, some of which were factual rather than subjective in nature. We also used a second approach to "smoking out" acquiescers, asking questions that demand a "no" answer. The results we obtained are presented in Table 8.5, A and B, and

Table 8.5A: Rates of Acquiescence in Response to Pairs
of Oppositely-Worded Questions

Questions	<u>Responding</u>	<u>Percent Acquiescing</u>
1. Are you usually happy?" vs. Are you usually sad?	23	27.8%
2. Are you usually with other people? vs. Are you usually alone?	25	56.0%
3. Most mornings here, do you pick out what clothes to wear? vs. Most mornings here, does somebody else tell you what clothes to wear?	21	71.4%
4. Do you live in G-10 (incorrect cottage number) right now? vs. Do you live in G- (client's cottage number) right now?	25	20.0%
5. Here is a picture of (superintendent's name). Is he the superintendent here at the colony? vs. Here is a picture of (business manager's name). Is he the superintendent here at the colony?	24	83.3%
Average:		55.7%

they must be seen to be believed. Unless answers to yes-no questions were viewed with a jaundiced eye, one might conclude that a large proportion of this sample consists of Chinese school bus drivers who fly airplanes on the side! It is important to note, however, that acquiescence rates varied from question to question. For example, although these respondents demonstrated an almost total lack of knowledge of the identity of the superintendent of the institution and acquiesced in response to the pair of questions about this topic over 80% of the time, they apparently were more certain of their cottage numbers and only 20% acquiesced when asked if they lived in another cottage.

To explore the implications of these staggering rates of acquiescence, we examined the extent to which these same respondents claimed to be involved in four chores in the cottage. The average respondent claimed to be involved in 80% of the chores; yet when we asked attendants the same questions, they claimed the average resident was involved in only 59% of the chores. More importantly, those respondents who showed themselves to be especially prone to acquiesce tended to claim more involvement in chores. If we had relied uncritically on these answers, we would have concluded, on the basis of a correlation of $-.61$, that lower IQ respondents were more involved in chores than

Table 8.5B: In Response to Questions for which
"No" is the Correct Answer

Question	Responding	% Yes Responses
1. Are you Chinese?	25	44.0%
2. Does it ever snow here in the summer?	26	73.1%
3. Do you know how to fly an airplane?	25	52.0%
4. Are you a school bus driver?	24	41.7%
5. Right now, is it raining outside? (or sunny outside, whichever question should be answered "no" at the time of the interview)	25	28.0%
Average:		47.8%

higher IQ respondents. Yet this finding would have been an artifact, for lower IQ respondents also scored higher on our measures of tendency to acquiesce. Using attendant data in place of client data, the correlation between IQ and involvement in chores turned out to be .26 rather than -.61. Thus this analysis not only confirmed the finding that acquiescence tends to be more likely among the more severely retarded but also suggested that their acquiescence can distort the relationship between IQ and the topics of interest in an interview (in this case, involvement in chores). We trust that these additional analyses leave no doubt about the magnitude of the problem of acquiescence. Responses to yes-no questions simply cannot be trusted.

The Use of Pictures With Either-Or Questions

In a continuing effort to maximize responsiveness while minimizing systematic bias, we tested some either-or questions which offered pictures as response alternatives, hoping to increase responsiveness to either-or questions without decreasing their generally high validity.

QUESTIONS ASKED OF COMMUNITY CHILDREN

Illustrative Comparison

Community children were asked whether they usually spend their time alone or with others, using the following set of four questions:

When you're not at school, are you usually by yourself or with other people?

When you're not at school, are you usually with other people or by yourself?

This picture shows a boy/girl who is by himself/herself and

this one shows a boy/girl who is with other people. When you're not at school, which boy/girl is most like you? Point to the picture.

This picture . . . (same as above with order of response options reversed)

Data collected from these four questions are summarized in Table 8.6.

As expected, the children were considerably more responsive to the two picture choice questions than to the verbal either-or questions. For the set of four questions, there are six unique pairs, allowing for six consistency figures. One consistency comparison is a within-format comparison between the two verbal either-or questions. The consistency figure for that comparison was only 58.1%. The largest fraction of inconsistencies for this comparison (25.6% of response pairs) occurred because the client chose the last option in both cases, suggesting a last option position preference.

A second consistency comparison was between the two picture choice questions. Consistency for this comparison was a respectable 72.7%. Inconsistencies were nearly equally balanced, with a slightly larger proportion of inconsistencies due to choice of the first option in both cases (14.7% of comparisons), suggesting that response bias may be lessened or possibly reversed in the picture choice format. The high consistency of these questions casts picture either-or questions in a favorable light.

The remaining four consistency comparisons were between picture choice and verbal either-or formats. These consistency figures ranged from 55.1% to 63.0%, averaging 59.3%. The patterns of inconsistencies among these questions were not particularly revealing.

Table 8.6: Either-Or vs. Either-Or with Pictures

Question	Responsiveness	Mean Consistency*	% Agreement
When you're not at school, are you usually by yourself or with other people?	69.6%	57.9%	51.0%
When you're not at school, are you usually with other people or by yourself?	80.0%	59.3%	50.0%
This picture shows a boy/girl who is by himself/herself and this one shows a boy/girl who is with other people. When you're not at school, which boy/girl is most like you? Point to the picture.	94.7%	65.4%	41.5%
This picture shows a boy/girl who is with other people, and this one shows a boy/girl who is by himself/herself. When you're not at school, which boy/girl is most like you? Point to the picture.	94.7%	62.2%	50.0%
Average:	84.8%	61.2%	48.1%

100

101

Average parent-child agreement for the four questions was only 48.1%, which was below the level of agreement expected if both parents and children selected their responses randomly. Further, the two formats differed little in terms of this estimate of validity. Agreement for the verbal either-or questions averaged 50.5%, as opposed to 45.8% for the picture choice questions. Thus, the major conclusion suggested by the agreement data is that neither format elicited substantial agreement with parents on this topic. Still, the findings point to the effectiveness of pictures as a supplement to these either-or questions, for they increased responsiveness and reduced a bias toward selecting the last option given.

Additional Comparison

A similar set of four questions, systematically varying both format and order of response options, was asked of community children. These questions asked if the child is usually happy or sad. Responsiveness to the two picture either-or questions averaged 97.4%, exceeding the responsiveness figure for the two verbal either-or questions (90.0%). Consistency for the comparison of the two verbal either-or questions was 75.5%. Most inconsistencies (15.1% of comparisons) involved choice of the first option in response to both questions, demonstrating that last-option bias for these questions is not a very robust phenomenon. Within-format consistency for the two picture choice questions was 67.8%, with inconsistencies more or less evenly mixed. For the four comparisons of picture and verbal either-or questions, consistency averaged 72.3%, with no pattern readily apparent in the inconsistencies. Average agreement between parents and children for the verbal questions

was 79.0%, as opposed to 70.4% for the picture either-or questions.

Conclusions

Both comparisons indicate that the use of pictures as response alternatives maximized responsiveness to either-or questions, a finding consistent with nearly all data previously analyzed. As regards the questions of reliability and validity, no clear conclusions can be drawn from these comparisons, which generated inconsistent results. The first comparison suggested a reliability advantage for the picture choice format, while the second comparison contradicted that result. Agreement data were also inconsistent. Agreement for the alone-with others comparison was so abysmally low that both either-or formats appear inadequate. For the happy-sad comparison, both formats appear in a favorable light, with verbal questions gaining a slight (79.0% vs. 70.4%) advantage. Overall, pictures clearly increase the ability of retarded persons to answer either-or questions. Since pictures do not appear to systematically threaten validity, their use can be recommended.

INTERVIEWS WITH SEVERELY RETARDED PERSONS

Alternative format comparisons to this point have left us trapped by a dilemma; namely, the very questions that many severely retarded persons can answer, particularly yes-no questions, yield invalid answers. It appeared that picture choice either-or questions offered a last hope of obtaining valid responses from retarded persons with extremely limited verbal skills. For this reason, we made a final attempt to test the effectiveness of picture either-or questions as a means of gaining information from severely retarded individuals (children and adults in the institution

who had been relatively unresponsive when they were first interviewed).

Illustrative Comparison

This comparison was identical to the happy-sad comparison conducted in the community children sample, pitting a verbal either-or question and its oppositely worded counterpart against a picture choice either-or question and its complement.

Responsiveness data for this comparison are represented in Table 8.7. Once again, the use of the pictures

as response alternatives maximized responsiveness. Comparison of the two verbal either-or questions revealed consistency of 71.5%. Moreover, this comparison suggested that last option response bias played an important role, as all inconsistencies (28.6% of comparisons) involved choice of the last option both times. Within-format consistency for the picture either-or questions was comparable (73.0%) but, importantly, gave no indication of systematic bias. Of the seven respondents who gave incon-

Table 8.7: Verbal and Picture Choice Either-Or Questions asked of Severely Retarded Persons

Question	Responsiveness	% Agreement
Are you usually happy or sad?	62.1%	63.6%
Are you usually sad or happy?	75.9%	96.0%
Format Average:	69.0%	79.8%
This picture shows a boy/girl who is happy, and this one shows a boy/girl who is sad. Which picture shows how you usually feel?	82.8%	73.1%
This picture shows a boy/girl who is happy, and this one shows a boy/girl who is sad. Which picture shows how you usually feel?	82.8%	66.7%
Format Average:	82.8%	69.9%
Overall Average:	75.9%	74.9%

sistent responses to the two questions, four did so by selecting the first option both times, as opposed to three who chose the last option both times. Thus, consistency data suggest that the two formats are comparable with regard to reliability, but that picture choice questions have an advantage in that they are not affected by strong response biases.

Agreement data are also summarized in Table 8.7. An apparent advantage exists for the verbal questions, as agreement for those questions averages 79.8% as compared with 69.9% for the picture questions. However, agreement figures for the two verbal questions were extremely discrepant. Although attendants consistently responded "happy," clients' responses changed as a function of the order of response alternatives. Thus, a high agreement figure (96.0%) for the question "Are you usually sad or happy?" represents not response validity but an accidental correspondence of client response bias with attendants' response preference. This is evident because agreement for the oppositely worded questions was only 63.6% and most disagreements involved choosing the last option ("sad") while attendants said "happy." By comparison, agreement for the two picture choice questions was relatively consistent and not powerfully affected by systematic bias; thus it can be concluded that the picture choice questions generate reasonably valid responses and are less distorted by response bias than the verbal choice questions.

Additional Comparisons

Two additional comparisons evaluated picture choice either-or questions used in interviews with severely retarded persons. The first comparison was identical in form to the above

comparison, pitting two pictorial against two verbal either-or questions, asking in each case whether the client is usually alone or with other people. As always, the advantage in eliciting appropriate responses accrued to the picture choice question, by a margin of 81.0% to 55.2%. Within-format consistency for the two verbal choice questions was 62.6%. Fully 25.0% of comparisons were inconsistencies in which the client chose the last option both times, indicating the importance of last option response bias in this severely retarded group. The within-format comparison for the picture choice questions yielded a higher consistency figure of 88.0%, with no evidence of order effects. Agreement for these questions was quite low, averaging only 58.3% for the verbal questions and 53.8% for the picture choice questions. Analysis of the inconsistencies revealed that they were in each case almost perfectly split between first and last choice responses by clients. On both verbal and pictorial questions, clients were much more inclined than attendants to respond "alone," implying that this is a genuine difference of opinion between clients and attendants which is not a matter of response set. Whether this disagreement involves invalid responding by clients, a lack of accurate information on the part of attendants, or is simply a real disagreement is impossible to determine.

The second comparison compared two picture choice either-or questions with three yes-no questions asking whether the client or someone else usually picks the client's clothes. The pictorial either-or clothes questions yielded an average responsiveness figure of 82.8%, which exceeded average responsiveness for the yes-no questions (67.8%). The comparison between the two picture-choice questions yielded a consistency figure of 62.5%. In

this case, 29.2% of the response pairings involved choice of the first option both times, suggesting that a first option response bias may sometimes be involved in picture choice questions. Analysis of agreement yielded results similar to those from the above comparison; agreement was only 52.9%, and in both questions clients were much more likely than attendants to assert that they chose their own clothes.

Conclusions

The one unequivocal conclusion which can be drawn from comparisons of verbal and picture choice either-or questions is that picture choice questions elicit appropriate responses from a larger fraction of the population. Many of these comparisons also suggest impressive reliability of responses to either-or questions, and particularly to picture choice either-or questions, which are sometimes less affected by response bias than are their verbal counterparts. These data indicate that the use of pictures does not reduce validity, although in some cases it introduces a first option response bias which parallels the last option bias found on verbal either-or questions. Overall, picture choice questions appear to hold promise greater than that of any other technique we have used as a method of gaining information from more severely retarded persons.

Conclusions

This chapter has examined the advantages and disadvantages of alternative approaches to obtaining information from mentally retarded persons. Viewed in conjunction with data presented in earlier chapters, it enables us to formulate a set of empirically-based recommendations to follow if one wants to obtain valid answers from mentally retarded interviewees.

Broadly, our findings establish that what one hears from mentally retarded persons depends in part on how questions are asked. We have demonstrated that not only question format but question wording and whether or not one uses probes after a question is asked can have dramatic effects on the answers that are given. As a result, the central recommendation of this study is as follows: *Whenever mentally retarded persons are to be interviewed, there must be careful attention to methodological issues, particularly question design.* The reader who has followed our narrative this far can scarcely be surprised and may even consider the broad recommendation so basic that it does not need stating. Yet, despite the fact that many studies relying on interviews with retarded persons have been conducted, this is the first time, to our knowledge, that this crucially important maxim has been stated. It is certainly the first time that measurement issues in interviewing retarded persons have been systematically studied. It is now time to end the tradition in the mental retardation field of simply assuming that responses to interviews are valid.

Based on the findings we have presented, we are able to offer the following more specific generalizations and recommendations.

1. Question wording matters. For example, if a question calling for an enumeration of activities in a given category (e.g., sports) mentions examples (e.g., football and baseball), retarded persons will be more likely than usual to name the very activities used as examples. Thus this strategy for making a question more concrete is counterproductive and should not be used. In addition, presumably small changes in question wording can have major effects on

response patterns. For instance, we found that retarded persons were far more likely to say certain unacceptable behaviors (e.g., hitting others) were prohibited if the question was phrased, "Are you allowed to. . ." than if the question was phrased "is it against the rules to. . ." We even found, thanks to a typographical error in the interview schedule, that responses to "Who decides what chores you do?" and "Who decides what chores to do?" differed. In addition, we found evidence that retarded persons do not always pick up distinctions in wording that the researchers hope they will; for example, distinctions among quantitative words such as "always," "usually," "sometimes," and "never." Finally, we found abundant evidence that responses to oppositely worded questions (e.g., "Are you usually with other people?" vs. "Are you usually alone?") can differ drastically. There are many examples that we can cite of misinterpretations, often humorous ones, of words and phrases in questions (see Sample 8.4). As we are not certain of all the possible effects of question wording, we can only recommend that there be extensive piloting of interview schedules to determine whether respondents understand and, if so, how they interpret alternatively phrased questions.

2. Techniques of probing for additional information must be carefully planned. As we have shown here, when interviewees are asked open-ended questions calling for an enumeration of activities, many cannot respond at all and those who can are likely to name very few activities. On the other hand, the strategy of probing indefinitely with "What else?" solves that problem but creates another that is more serious: overreporting of activities in response to the implied demand to do so to the extent that answers lose validity.

Thus, if such probes are used, and perhaps they should be to overcome the problem of underreporting on open-ended questions, they should probably be used only once or twice, not ad infinitum. Moreover, probes or follow-up questions should be phrased to hark back to the content of the original question (e.g., instead of "What else?" "What else do you do at your job?"). This strategy of reminding interviewees of the question content increases the ability of interviewees to answer, without reducing validity of response.

More generally, all of our findings on question wording and format need to be applied when probes or follow-up questions are used. Probes and follow-up questions should, as a general rule, be as carefully planned as the original questions they follow. Leaving the design of such questions to interviewers who are insensitive to the implications of question structure and wording is essentially leaving matters to chance. If one interviewer chooses to follow up an unclear answer with a simple yes-no questions while another chooses to use an either-or question to pursue the matter, we can predict that the two interviewers are likely to get different answers. This is by no means a trivial issue, for our experience suggests that no matter how well original questions are designed, there is likely to be some need to follow-up or clarify vague or confusing answers.

3. Multiple choice questions that call for responses along a quantitative dimension (e.g., a lot, some, not much, and not at all) appear to be generally useless as sources of information. They are relatively difficult for retarded persons to answer, in the first place; but more importantly, respondents appear to have difficulty attaching consistent meanings to such terms. We did not find any advantage in

Sample 8.4

A Sampling of Unusual Interpretations of Questions

1. Q: Which picture shows how you like the food here?

Point to the picture.

A: I know, keep your mouth closed.

2. Q: Do you decide what chores you do?

A: Probably clean house and fix your bed and clean up hallway
and mop and go walking and go see somebody's house.

3. Q: Are you allowed to hit people?

A: I don't hit people.

4. Q: What do you and your friends usually do together?

A: Let's see. My mom says... (Puts hand to head as if in
deep thought). Let me think. Well, I say yes.

5. Q: Do people here help you when you want help?

A: No, I help myself to bed, go to the bathroom.

6. Q: If you had one wish, what would you wish for?

A: Well, I go to school and we're going to have a party this
afternoon.

7. Q: Do you decide what chores you do?

A: We go to the canteen and get something.

8. Q: Do you decide what chores you do?

A: Not until I'm assigned to it.

9. Q: Do people here take things away from you?

A: No, but they do steal.

10. Q: Do you dust furniture?
A: No, wash ears, that's all.
11. Q: Are you allowed to go on dates here?
A: No, my daddy is.
12. Q: Has anybody from Voc. Rehab. ever tried to help you?
A: (nodded yes)
Q: How did they help you?
A: Wash your head.
13. Q: Where do you get your money?
A: My money, in my purse.
14. Q: Do people here make you do things you don't want to do?
A: It's not wise.
15. Q: Could you buy some candy with \$1.00 (question designed to check money concepts)
A: I done it, but I had pimples once before.
Q: (repeated)
A: If I don't get a hot dog.
16. Q: Do you usually have problems knowing what to do when you get sick?
A: No, always run to the bathroom and get sick.

breaking up a four-choice multiple choice question into a yes-no question and then a multiple choice question (e.g., "Do you leave here to go see your family?" if yes "Do you go see them a lot, sometimes, or not much?"). Nor did we find smiling and sad faces as an alternative to verbal multiple questions about satisfaction useful, except to boost responsiveness.

Unfortunately, we have no constructive advice for interviewers who seek quantitative information about such things as extent of involvement in various activities. Based on previous experience in other research projects with questions calling for estimated numbers of times doing something in a given time period, we steered away from that approach. We simply do not feel that retarded persons typically have a good enough grasp of quantitative concepts and time concepts that they can respond validly.

4. Multiple choice questions offering discrete response alternatives, however, may have promise. In asking respondents what type of dwelling they lived in and how they got to school, we found that questions of this type worked well; better than open-ended questions on the same topics, for they enhanced responsiveness without sacrificing validity. Although responses to verbal multiple choice questions were sometimes more affected by the order of options than were responses to parallel picture choice questions, we are not confident that pictures provide enough of an advantage to warrant the effort involved in producing them. Moreover, we are not able to say whether multiple choice questions naming discrete response options would work in other content areas, as we used them only to get information about basic, and presumably well-known, factual information.

5. Yes-no questions, despite the fact that they are easily answered by most retarded persons, are not useful in interviewing retarded persons. While this statement may seem strong, we have devoted more attention to yes-no questions than to any other format and we have found, with few exceptions, that retarded persons are highly likely to answer "yes" regardless of question content. Some of our evidence of acquiescence in response to yes-no questions cannot be described as anything other than shocking. In comparisons of oppositely worded yes-no questions, it was not uncommon for the most frequent pairing of responses to the two questions to be "yes" to both rather than the expected "yes" to one and "no" to the other. We have witnessed severely retarded persons, asked questions that should be answered "no," claim through their yeasaying, to be Chinese school bus drivers who fly planes on the side. The impact of acquiescence is so strong that we feel yes-no questions should be avoided entirely. If they are used, we can see no alternative but to build into the interview some check on acquiescence (see Chapter 9 for some suggested strategies).

6. Open-ended questions have one overriding limitation: they cannot be answered by very many retarded persons. As a result, they are not very useful at all in interviewing more severely retarded persons. However, they can be of use in interviewing persons with relatively advanced verbal skills. We found that simple open-ended questions in which a single response can be coded into categories (e.g., "Who decides what chores you do?") yielded relatively valid answers from those who could respond. Open-ended questions calling for enumerations (e.g., "What things do people here do that bother you?") were less satisfactory. In comparisons of them with parallel

sets of yes-no questions, we found that they probably led to underreporting, although it was not so pronounced as the overreporting in response to yes-no questions. Very few retarded persons appear to have the level of verbal skill that such enumeration requires. There is often no choice but to ask some open-ended questions in interviews. As long as the respondents are verbally skilled, these questions can yield great insight into the lives of retarded persons. The only caveat is that the insights gained will probably not be representative of the larger population of persons one is studying. Thus we can best envision open-ended questions being used as follow-up questions to more structured questions, to gain additional clarification of answers to those structured questions from those capable of providing it.

7. Of all the question formats we have tested, either-or questions, particularly those accompanied by pictorial representations of the alternatives, appear to be the best way of optimizing responsiveness and response validity. In numerous comparisons of either-or and yes-no questions, either-or questions consistently emerged victorious. While they are somewhat more difficult to answer than yes-no questions, this slight sacrifice in responsiveness is more than made up for by a gain in response validity. We have found some evidence of a preference for the last of the two options presented, but it has not been totally consistent and, more importantly, it has typically been a minor response bias in comparison to acquiescence. Furthermore, the use of pictures as adjuncts to either-or questions appears to be advantageous. It boosts responsiveness to the point that all or virtually all persons can respond and results in no major sacrifice of validity. If anything,

the use of pictures reduces the tendency to select the last option in either-or questions. This appeared to be particularly true among severely retarded persons, among whom response biases are generally more pronounced. Clearly either-or questions are not a panacea, for there are many topics that simply do not lend themselves to the either-or format. However, they should be given a great deal of consideration as viable alternatives to yes-no questions. For example, in a severely retarded group one could count on getting a huge number of "yes" answers to the questions, "Do you cook on a stove?" It would be possible to structure a pictorial either-or question on the same topic that would in all likelihood yield information that is much closer to the truth; for example, "Here is a boy who cooks on the stove. Here is a boy who does not cook on the stove. Which boy is like you? The one who cooks on the stove or the one who doesn't cook on the stove? Point to the picture."

In Chapter 9 we will spell out more of the implications of our findings for persons who solicit information from retarded persons by questioning them. As the list presented here suggests, our study generally revealed more about how not to ask questions of retarded persons than how to do so most effectively. Still, we believe that knowing what to avoid makes progress more likely.

CONCLUSIONS AND IMPLICATIONS

This report has covered a great deal of ground in detailing the findings of a complex project that was designed to explore the feasibility of interviewing as a means of obtaining information directly from mentally retarded persons. In an attempt to bring closure to the report, we will briefly review the study and its major findings, discuss recommendations for conducting survey research with mentally retarded persons, and point out some general implications for researchers and practitioners in the mental retardation field.

Summary of the Study

WHAT DID WE DO?

We set out to answer a question originally posed by the President's Committee on Mental Retardation: Is a national polling of a sample of the mentally retarded population a feasible means of gathering consumer input? We approached this guiding question by addressing the following more specific questions:

1. To what extent can retarded persons respond to questions in an appropriate fashion, and what factors affect their responsiveness?
2. How reliable are their responses, in the sense of being consistent over short periods of time?
3. How valid are their responses, in the sense of being free of systematic bias and agreeing with information provided by parents or caretakers or documented in records?
4. What types of questions appear to optimize responsiveness, reliability, and validity?

We began with reviews of the literature on the language and communication skills of retarded persons and response biases in survey research with the general population. We also explored issues in sampling and gaining access to retarded persons for purposes of interviewing them. Our literature reviews suggested that,

for the most part, retarded persons are likely to behave linguistically like children at various stages in the development of mature language competence, and that, moreover, many retarded persons have hearing and speech handicaps that may pose problems in an interview. In addition, we found cause in the literature on response biases in the general population to expect retarded interviewees, like children and less educated adults, to be especially prone to give biased answers; for example, to acquiesce in response to yes-no questions.

Our major mission then became one of systematically testing interviewing approaches with diverse samples of mentally retarded children and adults. Interview schedules focused on the topics of living circumstances, daily activities, utilization of services, and involvement in decision-making. Five samples of retarded persons were interviewed. The three primary samples consisted of 52 institutionalized children 12-16 years old, 58 institutionalized adults, and 57 children living in the community. Further research was done with 29 institutionalized and primarily severely retarded children and adults who had proven largely but not totally unresponsive in the first interviews, and a small sample of 13 adults living in the community and enrolled in a citizen advocacy program.

Each study also involved parallel interviews with "significant others": parents, attendants, or citizen advocates, depending on the sample. These persons were asked to answer questions paralleling most of the questions that had been asked of the retarded respondents so that the responses of retarded persons and nonretarded informants speaking for them could be compared. Finally, both institutionalized samples were interviewed twice, a week apart, so

that the reliability of their responses to the same questions asked on two different occasions could be examined. In all, well over 500 interviews were conducted. Alternative questions about the same topics were embedded within interview schedules to aid in identifying question formats and phrasings that optimize responsiveness, reliability, and validity of response. Much of the analysis focused on responses to individual questions, comparing answers given at different times, answers given to alternative questions on the same topics, and answers given by retarded persons and parallel answers given by parents or attendants. In addition, some composite scores were formed and used in correlational analyses so that individual differences in interview performance could be examined.

To highlight some of the major findings, we will offer concise answers to the four major questions of the study.

TO WHAT EXTENT CAN RETARDED PERSONS RESPOND TO INTERVIEW QUESTIONS?

In general, we found that many retarded persons are capable of giving appropriate answers to simply worded questions. The measure of responsiveness to questions that we typically used was the percentage of occasions on which a response met, at least minimally, the formal demands of the question (e.g., a "yes" or "no" to a yes-no question or mention of one activity in response to an open-ended question calling for an enumeration of activities). We obtained at least minimally appropriate responses at least two-thirds of the time in our samples of severely to mildly retarded persons. Of the various inappropriate responses we encountered, the most frequent appeared to be failure to give any response; vague answers that could not be coded into content categories

(e.g., equivocation on yes-no questions to the extent that the response could not be classified as "yes" or "no"), unintelligible responses, and irrelevant responses. An individual's ability to respond appropriately to questions was found to be highly stable from week to week, as indicated by correlations of .96 between percentages of answers that were appropriate in two interviews held a week apart with institutionalized children and adults.

Most importantly, we found it possible to predict how responsive individuals would be. First, responsiveness appears to be a function of IQ. As we had expected, it is all but impossible to interview most profoundly retarded persons. In our one attempt to interview profoundly retarded persons, only three of the 16 adults approached passed a simple screening interview and supplied any answers at all to the full interviews. Severely retarded persons were diverse in their ability to respond to questions, and the extent of their limitations compared to moderately and mildly retarded persons varied somewhat from sample to sample. Generally, however, responsiveness was greatest in the moderately and mildly retarded ranges. Overall, the correlations between IQ and responsiveness scores in the three major samples varied from .35 to .67.

Second, the ability to respond is a function of the type of question asked. Consistently in our samples, high numbers of interviewees were able to respond appropriately to simple yes-no questions and to questions involving choices among pictures in which a nonverbal response (pointing) would suffice. Either-or questions were the next easiest to answer, and open-ended questions and multiple choice questions with three or four structured options were the most difficult for interviewees to answer.

Finally, we found that responsiveness could be increased by simply repeating a question if the first response was less than minimally appropriate. In fact, some persons became more responsive to questions after having been interviewed with a similar instrument a week before. All of this suggests that retarded persons may need additional training and experience in being interviewed in order to learn how to respond appropriately, but that even during an interview they may benefit from additional chances to comprehend the question through having it repeated.

However, one of the major assumptions of the study was that obtaining answers is not enough. One must also determine whether answers given are reliable and valid.

HOW RELIABLE ARE ANSWERS?

In general, we can say that the answer given one week by a mentally retarded person will typically correspond to the answer to the same question given a week later. In both institutionalized samples, answers given a week apart were generally consistent over 80% of the time. This was certainly true of most yes-no, either-or, and open-ended questions. However, we found reliability very low for picture-choice questions about satisfaction with institution life and for multiple choice questions involving quantitative terms such as "a lot" and "sometimes." Finally, although we had expected answers to subjective questions to be relatively changeable, reliability figures for subjective yes-no questions were no lower than those for factual yes-no questions.

Overall, then, one can generally expect the picture of a group obtained one week to correspond to the picture of that group obtained the next week. One can expect retarded individuals to

"stick to their guns" in answering most of the time. We were not able to identify individual characteristics such as IQ that are closely related to the extent to which an individual responds reliably. Moreover, we encountered many problems in interpreting reliability figures. And finally, as any researcher knows, reliability is a necessary precondition of validity but is not sufficient to establish validity. Thus, we had to attempt to determine whether answers given were not only reliable but also valid and meaningful.

HOW VALID ARE ANSWERS?

The issue of response validity was addressed through two approaches. First, we compared responses of retarded persons to independent sources of validating information. Typically, this took the form of comparisons with the responses of attendants or parents to parallel questions, although some answers could also be compared to documented fact. Overall, retarded interviewees and attendants or parents agreed (gave matching responses) about two-thirds of the time in our various samples. At times, despite disagreement between individual pairs of clients and informants, pictures of the sample obtained from the two sources did not differ greatly. In such cases, one could at least use client and informant data interchangeably to describe a sample or to do analyses of individual differences. At other times, as a consequence of disagreement between pairs of clients and informants, pictures of the group obtained from the two sources were widely discrepant. As the overall finding of agreement about two-thirds of the time implies, it cannot be safely assumed that interviews with retarded persons will yield the same information as interviews with their caretakers or parents. This is true even when there is no reason

to suspect that clients' responses are systematically biased. It is particularly true when there is reason to believe that response biases such as acquiescence are at work.

Agreement with parents or attendants proved to be an imperfect index of response validity for two reasons. First, the two groups simply had different perspectives on some issues, and we were unwilling to infer that the retarded respondents were "wrong." Second, even high agreement was not always conclusive evidence for the validity of the responses of mentally retarded persons, for their answers sometimes revealed the workings of systematic bias. Thus, we took systematic biases in the response patterns of mentally retarded respondents into account in assessing response validity. We did so principally by comparing the responses of retarded persons to two or more questions on the same topic but with different wordings or formats. Where changing the way the questions were asked powerfully altered the content of responses, we inferred that systematic biases were reducing validity. Systematic rather than random discrepancies between the responses of retarded respondents and parents or attendants were interpreted in the same way. For example, we typically found clients far more likely to answer "yes" on yes-no questions than informants were, and this pattern indicated that acquiescence on the part of retarded persons represented a threat to validity of response.

In short, we combined agreement figures and evidence for systematic response biases in assessing response validity. In the case of yes-no questions, in spite of relatively high agreement figures, clear evidence of acquiescence led us to conclude that the responses of clients are largely invalid. Even when

agreement was high, we typically found more clients saying "yes" than attendants or parents. Moreover, in response to pairs of yes-no questions with reversed wording (e.g., "Are you usually happy?" vs. "Are you usually sad?"), retarded respondents frequently contradicted themselves by responding "yes" to both questions of the pair. Quite importantly, correlational analyses demonstrated that agreement percentages for questions increased as the percentage of parents or attendants saying "yes" increased; clear indication that agreement figures were often artificially high simply because acquiescence on the part of clients happened to put them in agreement with informants.

Open-ended questions provide another example of a question format for which high agreement was not necessarily good evidence for response validity. The vast majority of agreements involved both client and informant failing to mention something in a category of response. Somewhat more encouraging results were obtained with simpler open-ended questions that called for one answer rather than an enumeration of responses. Multiple choice questions with quantitative response alternatives (a lot, some, not much, etc.), as well as multiple choice questions calling for a selection of happy or sad faces to indicate a degree of satisfaction, appeared to have very little validity at all. Finally, either-or questions, as well as factual multiple choice questions with discrete rather than quantitative options, appeared to have relatively good validity. Despite some evidence of a tendency to prefer the last option in such questions, there typically was not a marked bias in the inconsistencies that occurred. Thus, we emerged from this analysis most confident about the validity of responses to open-ended questions

calling for a single answer, either-or questions, and multiple choice questions with discrete response alternatives.

We should note that analysis of responses to questions whose answers were documented fact did little to increase our confidence in answers given by retarded persons. Discouragingly high proportions of respondents were not able to supply fully correct answers to questions about their full names, birthdates, addresses, and so on. Many others could not respond at all to such questions. Analysis of these questions bolstered the conclusion that the validity of information provided by retarded persons can never be taken for granted. Moreover, it suggests the practical need for training retarded persons so that, particularly in emergencies, they can communicate basic facts about themselves accurately.

Finally, we were generally unable to predict which persons would tend to give valid answers and which would not. Surprisingly, children living in the community were no more likely to agree with their parents (at least on a set of yes-no questions) than institutionalized persons were to agree with their attendants. Persons with higher IQs were more likely to agree with their parents or attendants than persons with lower IQs, but the correlation was not strong enough that we could confidently predict whose answers can be trusted and whose cannot. Until further research is conducted, one must be wary of responses given by any retarded person, especially on the types of questions that are most likely to elicit systematic response biases.

WHICH QUESTIONING APPROACHES ARE MOST USEFUL?

A major undertaking in this project was to conduct head-to-head

competitions between alternative question formats and phrasings in order to determine, with the topic of questioning controlled, which alternative approach optimized responsiveness and response validity. To examine response validity, we again compared answers given by retarded persons and either attendants or parents, and we also looked for signs of systematic bias when the answers of clients to alternative questions were compared. Since Chapter 8 included a rather detailed summary of the findings of these analyses, we will only briefly review the recommendations that were made.

1. Careful attention should be devoted to question wording. Seemingly small changes in wording can have large effects on responses. For instance, the use of examples of the topic of inquiry in the question (e.g., mention of football and baseball as examples of sports to be enumerated) biases responses such that the examples are mentioned more often than they would otherwise be.
2. In probing for additional information on open-ended questions by asking "What else?" the probing should not be carried on indefinitely until the respondent can think of no more, for this strategy appears to threaten validity of response. It is probably wiser to use such probes only once or twice. Moreover, when probes are used, they should restate the content of the original question to increase understanding (e.g., "What else do you usually do for fun?" rather than "What else?").
3. Multiple choice questions with quantitative response options which are designed to get at
- extents of involvement in various activities appear to be generally useless, for retarded persons appear to have difficulty attaching consistent meanings to such terms as "a lot," and "not much." Similarly, pictures of happy and sad faces do not aid in eliciting information about satisfaction with life. However, multiple choice questions about basic facts that offer discrete response alternatives (e.g., walking, riding the bus, riding a bike, and riding in a car as ways of getting to school) do appear to be useful.
4. Yes-no questions should probably be avoided in interviewing retarded persons because of the often disastrous effects of acquiescence on response validity. Pictures as accompaniments to such questions apparently do nothing to prevent acquiescence.
5. Open-ended questions have the major disadvantage of not being answerable by high proportions of mentally retarded persons. Our tests of questions calling for enumerations of activities suggested that these questions yield very little information from most people and probably result in underreporting of the activities inquired about. By contrast, open-ended questions calling for single answers generally yield valid responses, assuming they can be answered, and appear to be appropriate in interviewing higher IQ retarded persons.
6. Either-or questions emerged as the single most satisfactory type of questioning strategy. Although they are not quite as easy to answer as yes-no questions (indeed, several respon-

dents answered them with a "yes"), responsiveness to them was still relatively good, and more importantly, clients' answers and informants' answers tended to be in agreement. The only negative finding was a tendency, but not a strong or totally consistent one, toward selection of the last of the two options presented, a response bias that would need to be guarded against. Finally, it was in conjunction with either-or questions that pictures appeared to have their greatest effect. The use of pictures with either-or questions consistently increased responsiveness to the questions without sacrificing validity (and in a few cases enhancing it by counteracting the bias in favor of the last response option).

We are painfully aware of the fact that our research has typically told us a great deal more about how not to interview retarded persons than how to do so effectively. As the list above suggests, we continually encountered threats to response validity that forced us to raise grave questions about the utility of some questioning tactics. Even for the best of question formats, validity was often marginal. Nonetheless, the findings do offer guidelines for constructing interview schedules to optimize the chances of getting answers, and more importantly, getting answers that are meaningful reflections of the attitudes, needs, and circumstances of retarded citizens.

DIRECTIONS FOR FUTURE RESEARCH

The present study breaks new ground by exploring in a systematic way methodological issues in interviewing retarded persons. Obviously the answers we have been able to provide are less than complete; more-

over, our study has raised at least as many questions as it has answered. It has at least established unambiguously the fact that the validity of answers given by retarded persons can never be taken for granted. However, many of the specific factors which influence the responses of mentally retarded persons remain unexplored. We would encourage other researchers in the mental retardation field to continue this promising line of research in hopes of enabling both researchers and practitioners to obtain useful information from retarded persons. We recommend careful investigation of the entire range of factors that may affect responses.

Our study has focused almost exclusively on characteristics of interviews that affect responses, primarily question format and phrasing. Other issues in interview design remain to be addressed. For example, our study suggested that social desirability may enter into the responses of retarded persons to many questions, but we did not systematically deal with that issue. Do mentally retarded persons have a heightened motivation to give socially approved answers, and are their responses especially biased on items that raise issues of self-presentation? A careful investigation of question difficulty, perhaps emphasizing vocabulary difficulty and conceptual complexity, might also prove fruitful, as this was only incidentally examined in the present study.

Moreover, as Sudman and Bradburn (1974) indicate, respondent variables and interviewer variables, along with task variables such as question design, can introduce error into survey research. We have explored the respondent variable of IQ in some detail, and have given some attention to other respondent variables such as age, sex, and institutionalization. Although we found

these additional respondent variables to be of minor importance, more intensive investigation of them might be warranted. The fact that we have found lower IQ individuals less responsive, more likely to acquiesce, and somewhat less likely to agree with parents and attendants emphasizes the potential importance of respondent variables. As for interviewer variables, we left the issue entirely unexplored. For example, do interviewers who are already familiar to retarded interviewees obtain different information than do outsiders? How do the interviewer's attitudes and behaviors affect responsiveness and the nature of responses? The literature on threats to response validity in survey research with the general population (See Chapter 2) provides a ready source of leads.

Meanwhile, researchers who are not directly interested in methodological issues in interviewing cannot afford to ignore them if they are conducting research with retarded persons that entails asking questions. Again, the literature on survey research, as well as the present monograph, can provide guidelines to follow in constructing instruments. The researcher sensitive to the issues will probably want to build checks on response validity into whatever instruments are developed. We are confident that researchers will detect many other leads in this monograph that we have not mentioned here.

How to Interview Mentally Retarded Persons

Based on our experience as well as our data, we would make the following recommendations to persons seeking information from mentally retarded individuals through survey research. (The reader is also referred back to Chapter 4 for more about the logistics of interviewing.)

PREPARATION

Invest a substantial amount of time in preparation before any "real" interviews are conducted.

It is not unusual in the mental retardation field or in any other area of research for researchers to "slap together" an interview or questionnaire. There is not a strong tradition in survey research, as there is in other areas of measurement such as testing, of establishing the reliability and validity of instruments. We have "slapped together" interview schedules for mentally retarded persons ourselves. While we pilot tested many questions used in the present study, our efforts could still be characterized as informal at best. Granted, we labored individually and in work groups for hours deciding what we wanted to ask and how to phrase and structure questions on those topics. We went through many draft versions of these interviews before we were satisfied. But, most of our planning was based on armchair thinking, on an attempt to imagine all the problems a question might present to a mentally retarded person. There were still numerous occasions in which we found, while administering interviews, that we had not anticipated everything, that we were getting bad information because our question was bad. To give just one example, we found that institutionalized persons often had difficulty with questions about their activities due to the fact that they were not sure whether we meant to ask them about what they did at the institution, what they did at home, or both. A few interviewees explicitly asked which we meant, but most others simply answered, and there is no way to tell how they interpreted the questions. If we had wanted to know their extent of involvement in, for example, home-making activities in the cottage at the institution, we would have had

to add the phrase "in your cottage" or the like to each question. We indeed did just that once we recognized that our initial questions were ambiguous, but by then we had collected data from whole samples and had no way to be sure whether their answers pertained to activities at the institution, at home, or both.

We would recommend starting with a serious "armchair" process of designing what appear to be simple and unambiguous questions about whatever topics are of interest. However, we would then recommend moving from the armchair to the field to do extensive piloting of the draft interview schedules with real retarded persons similar to those to be interviewed in the full-fledged study. No matter how well one anticipate exceptions, misinterpretations, and ambiguities, one is certain to detect more during the piloting process. If it is feasible, we would recommend conducting parallel interviews with attendants, parents or houseparents, or obtaining validating data from files and records so that the pilot testing process yields information about the validity of answers, beyond what can be obtained by seeing how interviewees are interpreting questions. Because interviewing retarded persons is particularly difficult, this kind of preparatory work appears to be crucial.

QUESTION DESIGN

Draw on the successes and failures of the present study in designing questions.

To be as brief as possible here, that would include paying careful attention to question wording and using question formats that are likely to optimize responsiveness and validity of response (see Chapter 8 for more details). We would recommend an interview schedule that is highly structured in basic design,

with open-ended questions used as follow-up questions in areas where more clarification is desired from more verbal respondents. We would recommend heavy use of either-or questions and multiple choice questions with discrete rather than quantitative options. If severely retarded persons are to be interviewed, such questions should, when possible, be accompanied by pictures. Simple open-ended questions calling for one piece of information rather than an enumeration would also be appropriate, although it must be recognized that many less verbal persons will not be able to answer them, even if they concern very basic facts. By using structured questions as the basis for the interview schedule, one can hope to obtain at least some information from virtually all persons in the sample. This information can be supplemented with elaborative and illustrative material gained from open-ended questions that only some respondents can answer. We would also recommend asking as few questions as possible that require concepts of number, time, quantity, monetary value, history, and the like.

Based on piloting, one can anticipate problems that interviewees are likely to have and design, in advance, paraphrases or alternative phrasing of questions to be used if interviewees either cannot understand or ask for clarification of the original question. Alternatively, the decision can be made to ask questions only in one standard phrasing and simply live with any missing data that result. The advantage of the first alternative is that one can hope to gather information from a larger portion of the sample under study; the risk is that the alternative way of asking for the information may change responses. The more like the original question the clarifying question is, the less that risk will be. Indeed, our analysis

of responsiveness suggests that simply repeating the question in the same form will increase responsiveness, with no alteration of the meaning of the question. Thus, such repetitions are clearly recommended. Still, there are inevitably some who cannot understand the question even when it is repeated.

Similarly, probes and follow-up questions generally should be planned in advance. A sensitive interviewer can do wonders in pursuing an answer and drawing a person out; as a result, we have a reluctance to achieve total standardization. However, our concern with the potential impacts of question format and phrasing on responses makes it necessary. Our recommendation that interview schedules be layered, with structured questions followed up by open-ended ones, helps in this regard. The structured questions would provide exactly the same stimuli to all respondents. Then to gain further detail from more verbal subjects, the interviewer might be given discretion to pursue interesting responses, as long as that interviewer is sensitive to the effects of question format and wording and will not lead the respondent.

We have recommended that yes-no questions be avoided if at all possible, and will stick with that recommendation. However, if they are used, we would strongly urge that checks for the operation of acquiescence be built into the interviewing process. Our data on acquiescence have suggested that it is not just a matter of isolating a few acquiescers; we have found instead that most retarded persons will acquiesce to some extent at some time. Still, the phenomenon is more pronounced for some persons. One possible check is to scatter oppositely worded questions of the type we have used throughout the interview schedule and determine how many times respon-

dents contradict themselves by answering "yes" to both questions in a pair. The other strategy we used, asking questions that should properly be answered "no," appears to be equally appropriate. One might wish to develop a formal acquiescence scale using one or both of these approaches, testing it out in advance to make sure it is reliable and valid. Even an informal check, however, would permit detection of frequent acquiescers. We are not yet in a position to suggest an explicit criterion for defining frequent acquiescence. It would be up to the researchers involved to establish such a criterion and then use it to eliminate acquiescers from the usable sample. However, because checks for acquiescence would inevitably mean disqualifying some responses, we continue to think that it is best to avoid asking yes-no questions if at all possible.

HEARING FROM SEVERELY AND PROFOUNDLY RETARDED PERSONS

Use alternative ways of obtaining information about and from more severely retarded persons.

While we have not always found strong relationships between IQ and the interview measures we have examined here, we have generally found more problems interviewing persons in lower IQ ranges. Our one attempt to interview profoundly retarded persons, done simply to document the obvious, was an almost total failure. Severely retarded persons, while less predictable, also have relative difficulty answering verbal questions. For them, we found the pictorial either-or format particularly promising, but we must still conclude that they have limited capacities to respond and that they are particularly susceptible to response biases such as acquiescence and a preference for the last option in an either-or question. The

two-layered interview schedule approach we have advocated here would perhaps allow one to obtain answers to structured questions from most severely retarded persons, but few of them would be able to elaborate on their answers in response to open-ended follow-up questions. Thus, there appears to be no way to avoid underrepresenting the perspectives of more severely retarded persons in survey research.

There are alternatives. One which has been used before is to obtain the desired information from a parent, attendant, houseparent, or other representative of the retarded person. However, there are limitations to this approach. We have at times encountered cases in which attendants or parents and retarded respondents simply disagreed about something, where one cannot explain discrepancy between their answers as a matter of response bias on the part of the retarded respondents. We have also found numerous instances in which attendants and parents have surprised us by not knowing something that we assumed they would know about the retarded persons they interact with.

On the other hand, we feel that relying on such persons as informants is better than simply writing off severely and profoundly retarded persons. Indeed, even when mildly and moderately retarded persons are interviewed, we would recommend interviewing their "significant others" as well, partly as a check on the validity of certain responses, and partly to obtain information that most retarded persons have difficulty providing themselves (e.g., information about history, involvement with service agencies, financial matters, and amounts of time involved in various activities). Parents and care providers are especially valuable as sources of factual information; whether they can provide

information about retarded persons' perspectives and opinions is another matter. We would urge that severely retarded persons still be given an opportunity to speak for themselves, particularly about their opinions, even when nonretarded informants must be the primary source of information about them.

Another alternative to interviewing low verbal individuals is behavioral observation. Instead of asking them about daily activities, one might observe them through time sampling in the residential setting. Instead of asking about satisfaction with the residential setting or with service providers, one could potentially observe affective responses to service providers or to events in the residential setting. We are not claiming that this would be easy, for behavioral observation is time-consuming and difficult. However, that might be the best way, coupled with interviews of informants knowledgeable about the retarded person, to approximate an understanding of the needs, perspectives, and circumstances of persons with limited or non-existent verbal skills.

To be sure, the data would not be comparable to data obtained from more verbal individuals through interviews. Thus, we have no easy solution to the problem of obtaining comparable data from the entire range of retarded persons. When previous researchers (e.g., Bell, 1976) have relied on informants to answer for low-IQ persons but have obtained answers directly from higher-IQ persons, there is simply no way to tell whether any resulting differences between high and low-IQ persons are due to real differences in their lives or due to the difference in data sources. This same quandary would be true any time the source of information for one group is different from the source of information for another. However, it

would be possible to interview both clients and nonretarded informants in the high IQ range, and use the agreement between their responses as a way of roughly estimating the degree to which responses given by nonretarded informants and severely retarded persons might agree. If agreement is high, one would feel more comfortable combining nonretarded informants' data for low-IQ persons with retarded informants' data for high-IQ ones.

We should perhaps make some mention of the nonverbal communication systems now used with language-deficient retarded persons. In our opinion, while such systems are permitting great advances in the communication training of severely and profoundly retarded persons, the time has not yet arrived when they can be used to collect data from samples of retarded persons on a widespread basis. The major problem is that even if interviewers could be trained to administer the interview in a nonverbal communication system, what nonverbal communication system would be used? There is no standardization at this point, as there would be if one wanted to interview deaf persons and could use Ameslan as the vehicle. While Ameslan has been used with some retarded persons, many other systems, including sign languages developed especially for use within a particular facility, are too provincial to be used in a large-scale interviewing project. It might be possible to enlist the help of facility staff to "translate" interview questions into a local nonverbal communication system. However, one would then have to guard against the possibility that respondents would react differently to a staff interviewer than to an outside interviewer. All things considered, then, these systems do not offer much hope at the present time, except, of course,

as a means for facility staff to gather input from their own residents for program planning purposes.

Implications of the Study for the Mental Retardation Field

THE FEASIBILITY OF A NATIONAL POLLING

This project was inspired by the President's Committee on Mental Retardation, which wanted to know whether or not a periodic national polling of retarded citizens would be feasible as a means of obtaining consumer input. In light of our findings, the answer to this question depends on the form this proposed polling takes. A comprehensive survey of mentally retarded consumers, comprehensive in the sense of involving the whole range of mentally retarded persons, and comprehensive in the sense of covering a broad spectrum of topics relevant to the lives of the retarded, appears not to be feasible. Very simply, too many retarded persons would have too much difficulty providing valid information about too many topics. However, less inclusive surveys, constructed and administered with our findings in mind, could provide PCMR and others with valuable information about the needs, circumstances, and attitudes of particular groups of retarded consumers.

The data suggest that a standardized, comprehensive, national survey of mentally retarded persons is impractical on two major grounds. First, it is clear that standardized interviewing of the entire spectrum of mentally retarded consumers is not possible. Our experience indicates that interviewing profoundly retarded persons is clearly impracticable, and that the population of severely retarded consumers includes many people who are unresponsive or provide responses of very low validity. Further, even among the mildly and moderately retarded, many persons

give responses to interview questions which are largely invalid. Serious attempts to interview retarded persons must include validity checks which screen out the answers of people who consistently display response biases. Thus, survey research with retarded consumers necessarily involves a selective sample of more verbally proficient and intelligent persons, and the information gained might be quite unrepresentative of the entire national population.

Secondly, the data presented in the preceding chapters indicate that there are serious restrictions on the types of information which can be obtained from interviewers with mentally retarded respondents. The only question formats which generated information of reasonable validity in our project were either-or questions, multiple choice questions offering discrete rather than quantitative response options, and open-ended questions targeted at discrete bits of factual information. Restriction of an interview schedule to these formats limits the types of information that can be sought. Further, we have found that certain types of content areas are especially problematic; specifically, questions using such concepts as time, number, quantity, monetary value, history, and so on are difficult for mentally retarded persons and tend not to generate meaningful answers. Thus, we cannot conduct a polling of retarded consumers which is comprehensive with regard to topics included.

Our study indicated unambiguously that retarded persons would have too much difficulty providing valid information for a comprehensive national survey to be feasible. We have not discussed some of the formidable problems in constituting a national sample; distinguishing between the identified population and the "actual" population would be difficult, and

gaining access to even the identified population, given current safeguards protecting the confidentiality of client records, would be virtually impossible.

However, even if a comprehensive national survey appears to be a poor means of information gathering, the goal of hearing from retarded individuals remains an important and viable one. We would be enormously discouraged if our study caused people who work with the retarded to give up; to stop trying to hear from their clients regarding crucial issues in their lives. Rather, implicit in our statement of what we cannot do is a host of information regarding what we can do, and it is to the possibilities for communicating with mentally retarded consumers that we turn next.

THE NEED FOR PERSISTENCE

Even though we have had, at times, to be pessimistic about interviewing retarded persons, our data do indicate that one can interview some of the people some of the time. It is our hope that this monograph, rather than discouraging people from attempting to gather input directly from mentally retarded consumers, will encourage them to do so in a methodologically sound way. As we see it, the most serious problems we have encountered are those that involve systematic response biases. Acquiescence is the single most serious problem we have encountered, and efforts to guard against it must certainly be made. However, when there is no systematic response effect operating, the threat to validity is usually not as great. Granted, there still may be error in the data of substantial size, but the picture one obtains of a group through one questioning technique may not differ greatly from the picture one obtains from an alternative technique. At other times, we have found retarded persons able to

answer and able to answer with valid information. Furthermore, we have suggested some ways of optimizing responsiveness and response validity which we hope will be useful to others. We have assumed and continue to assume that because communication is a two-way street, communication with retarded persons can be improved if interviewers improve their techniques of soliciting information.

However, our experience with retarded persons has instilled in us a commitment more basic than our commitment to increased methodological sophistication. That commitment is to an increased sense of humanity and sensitivity in our dealings with mentally retarded persons. Our empirical data somehow miss a vital point about our interactions with the retarded: namely, that nearly every one of our respondents, even low verbal individuals, had something to say. Repeatedly, people we interviewed told or showed us that being interviewed was for them a rare, exciting, and gratifying experience. Perhaps this is a commentary on the fact that their views are often ignored, or at least that nonretarded people rarely sit down for any length of time with them individually and ask them to talk about themselves. For the interviewer who enters into the task with a genuine regard for retarded people and a real desire to learn about them, the interviewing process is equally rewarding. In fact, we developed a strong sense that in those interviews which did not elicit much information, the failure resulted not because the mentally retarded interviewees did not have the right answers, but because we *did not ask the right questions*. Our structured interview schedules not only taxed the communication skills of our interviewees, but often failed to give them a genuine opportunity to speak their minds. A more flexible

approach, tailored to the capabilities and interests of the individual client, frequently produces a less rigorous but richer interaction. At least we found this to be the case in our more free-wheeling interviews with adults living in the community.

Thus, in its entirety, our study provides a set of more or less specific recommendations for the field of mental retardation. First, the idea of a comprehensive national polling of retarded consumers is probably not practical, at least given our current knowledge. However, a segment of the retarded population can respond to traditional, standardized interview techniques, and can give valid information if interviewed within the constraints discussed in the previous section. We suggest that workers in the field intensify efforts to interview, with methodological rigor, particular groups of mentally retarded consumers of services. In addition, even when formal interviewing is not possible, given the person or the topic, retarded persons may still have a great deal to say. For those individuals and topic areas not amenable to formal, standardized interviews, we recommend careful efforts to communicate in less structured, less formal interviews, tailored to the needs and capabilities of the individual client. Our project generated some useful guidelines as to how this process of informal communication can be made more efficient and can generate more valid information. It is to the guidelines that we can turn next.

ISSUES FOR CONSIDERATION BY PRACTITIONERS

For the most part, this monograph has been directed toward researchers and others who conduct formal interviews with retarded persons for data-gathering purposes. Yet there is no great difference,

except for degree of formality, between a survey research project and the kinds of informal questioning and interviewing that case-workers, counselors, and other professionals do daily with retarded clients. All major legislation affecting retarded persons today, the Education for All Handicapped Children Act and the 1978 rehabilitation amendments, including the developmental disabilities sections, calls for the development of individual program plans. Moreover, they call for the involvement of clients or students, or their representatives, in the development of such plans. Beyond the individual program planning process, professionals who work with the retarded frequently need to obtain information from them and to interact with them verbally in order to carry out treatment and training efforts and to evaluate individual progress and overall programs.

We view our findings as just as relevant to the practitioner as to the researcher. Take, for example, the matter of yes-no questions. Everyone who talks to a young child or a retarded person with limited verbal skills slips into using them. Yet our findings clearly indicate that retarded persons are highly likely to say "yes" to such questions, whether "yes" is the appropriate answer or not, and that the answers one obtains to two yes-no questions that are oppositely worded are quite likely to be discrepant as a result. We think practitioners should be alert to the possibility of acquiescence in their discussions with retarded clients, particularly if important decisions are going to be based on whether the client says "yes" or "no." One might want to consciously attempt to re-ask a yes-no question later in the interview with the wording reversed (for example, after asking, "Do you think

the car wash would be a good place to work?" ask, "Do you think the car wash would be a bad place to work?"). If the answers are both "yes," one at least knows that further questioning will be necessary to clarify what the client really does want. Alternatively, as we have suggested here, an either-or question ("Do you think the car wash is a good place to work or a bad place to work?") might be a better way of obtaining a valid answer.

The practitioner might benefit in other ways from being more sensitive to the effects of question wording and format. Use of the simplest wordings possible is a starting point. Again, it is useful to ask the same question in two or three ways to check on client comprehension. Wording is particularly problematic when one is asking about service agencies. Our studies, for example, have suggested that clients have varied understanding of what vocational rehabilitation is. If asked whether they have been a vocational rehabilitation client, retarded persons may not understand that wording, but may recognize "voc rehab," or "TRC," or another local name for the rehabilitation agency. (We have some evidence to suggest that they may be even more likely to talk about their experiences with an agency if you are able to name the counselor they actually had. Many of our respondents related to a service agency in terms of a specific staff member, sometimes not being able to identify the agency or what its functions as a whole was called or was all about.)

As we have also suggested, using examples to make a question more concrete may backfire by changing the meaning of the question or by prompting mention of the examples in answers. Generally, then, one would want to be careful not to lead the

client toward a response that he or she might otherwise not give.

When the same information is being sought from several clients (for example, in a program evaluation survey), the same standardization of questions that we would recommend for researchers should be used. One often cannot predict how much a slight change in wording might affect the answers that are given. On the other hand, we are quite confident that if some clients are asked a yes-no question and others are asked an either-or question on the same topic, their answers are likely to differ. This clearly points to the need for careful planning in designing the questions to be used and adherence to those questions in actually conducting the interviews. The interviewer might still be given latitude to pursue interesting or confusing answers to the standard questions. To be sure, the practitioner does more informal types of data-gathering than the researcher, and does not have as much concern with methodological issues. But this does not mean that the practitioner does not need to be concerned with the validity of the answers that are given.

Finally, a note to teachers and trainers of the mentally retarded is in order. Our findings are highly relevant to teaching communication skills to retarded persons. Overall, they suggest that many retarded persons are in need of training which would better equip them to understand and answer questions, not only in interview situations but in everyday exchanges with others. Too often, we feel, language and communication training emphasizes the formal rules of language but does not go far enough to help retarded students use language appropriately in different social contexts. Retarded persons need explicit training in answering

questions appropriately so that their answers meet the demands of the questions (e.g., so that they answer an either-or question with one of the two options rather than by saying "yes" as many of our subjects did). Beyond that, as our findings clearly suggest, they may need guidance in answering questions accurately, especially questions about basic personal information.

More specifically, we would offer the following suggestions for inclusion in communication training programs. First, concentrate on teaching retarded persons to answer questions appropriately and to avoid answers that are unintelligible, irrelevant to the question, or so vague that they are uninformative. Our data on responsiveness to questions suggests a developmental hierarchy of types of questions, progressing from yes-no questions and questions involving pictures that can be answered by pointing, to either-or questions, to multiple choice questions, and finally to open-ended questions. Although we have not emphasized it here, there are also levels of difficulty within the category of open-ended questions. For example, in previous research (Sigelman & Werder, 1975), retarded adults were more responsive to what-questions asking them to name or describe something than they were to why-questions and questions that asked them to consider hypothetical situations (e.g., "What would make things better for you where you live?"). Training might involve asking students progressively more difficult types of questions while giving them feedback and instruction as to what the question is demanding. Once the student can supply minimally appropriate responses, the trainer might encourage them to provide more complete responses and elaborate on their answers. For instance, if the question is "Do you have a job?" a

minimal response might be "yes," a better response might be "Yes, I have a job," and an even better one might be "Yes, I work part-time at a cafeteria."

In addition, training should be directed toward obtaining valid answers, not just answers. Two points are particularly worth emphasizing. First, as we have shown, many retarded persons are not able to answer basic questions about themselves such as full name, birthdate, address, and so on. For some, the problem is answering an open-ended question at all, due to limited verbal skills. However, we feel that even those with limited verbal skills could be taught to respond by rote to a set of basic questions about themselves. For others, the difficulty lies not in answering but in answering accurately, and these individuals also need training. We feel strongly that retarded persons should be able (for example, in an emergency) to tell people such things as who they are, where they live, their age or birthday, their telephone number, and so on.

Secondly, we feel that acquiescence and possibly other response biases should be given attention in communication training programs. One simple way to do this is to have students practice answering questions about concrete objects and materials or about verifiable facts in their immediate environments (e.g., "Is your shirt blue?" "Am I your father?"). This allows the instructor to provide feedback about the accuracy of answers that are given as well as their form. Quite obviously, retarded persons can get into trouble in the "real world" if they continually say "yes" when they are asked questions calling for a "yes" or "no" answer. Rosen, Clark and Kivitz (1977) have explored this acquiescence problem in the larger context of compliance with requests and have found,

for example, that many retarded persons will agree to take a pill not knowing anything about what the pill contains. In every area of their existence, from sex to individual program planning, retarded persons must learn to avoid exploitation and assert themselves. At least part of this larger mission can be accomplished by teaching them not to say "yes" without thinking about the answer.

CONCLUDING REMARKS

We would like to leave our readers with a final message: retarded people must finally be treated as people. For service providers and researchers alike, this concern with basic rights includes an obligation to obtain information concerning the perspectives of the retarded. Too often we have not bothered to ask retarded people themselves what is going on and what they think of it. We sometimes assume too readily that a residential program or an educational activity or a service system that looks good to us must look good as well to the retarded clients it serves. We have relied too often on the opinions of "significant others" to determine what retarded individuals need or want. Although data obtained directly from retarded persons through interviews cannot totally replace other sources of information, it should be one of the ways in which we attempt to "know" retarded persons. If questions are hastily designed without attention to measurement issues of the kind discussed here, the gains will be few. Indeed, the data collected might easily be misleading; for example, if program evaluators ask residents of a facility questions like "Are you satisfied with the training here?" they thereby assure themselves of a favorable evaluation report. However, if care is taken in deciding how to seek information, the gains can be great. Furthermore, our data clarify

an important point: even informal interactions with mentally retarded persons must be conducted with sensitivity to the level of their communication skills. A well-intentioned effort to "make things easier" for a mentally retarded person by phrasing inquiries as yes-no questions can have the effect of destroying that person's opportunity to communicate.

In closing, we cite two of the less-structured interviews we conducted with mildly retarded, relatively verbal adults living in the community, in which we asked them to reflect on the implications of being retarded. (Virtually all of these adults denied being retarded, but had much to say about the implications of being considered retarded.) We believe these transcripts are an excellent illustration of the value of asking retarded persons to speak for themselves.

Interview I

Q: What bad things happen to you when people think you are retarded?

A: Not too much has happened to me bad. They been pickin on me. They say, "You can't do that because you're mentally retarded. You can't do this." They just sort of push me back when I want to do something and say, "Oh, you can't come with us. We don't want you." They just push you back into a corner. "She can't do this! She doesn't know what she's doing."

Q: What good things happen to you when people think you're retarded?

A: Oh, they seem to want to take me out to different places like football stadiums and picnics and going out to the falls and spending the day out there. And going on all day trips.

Q: Does anything else good happen?

A: No.

Q: O.K., what would it be like for you if people stopped thinking you were retarded?

A: It'd be pretty good. We could probably get along a lot better if people didn't think of us as being retarded and handicapped. We'd probably make it out on our own. But see, the rest of the people are just pushing you like in one certain square and saying, "That's where you can be; you can't come over here." Because I've had that all my life.

Q: How do you think it would be different? What would you be doing differently if people hadn't done that to you?

A: Uh, I think I'd have higher paying jobs and be able to go to restricted movies; not that we want to, but you know, if you wanted to, be able to do it and not having to say, "We don't allow you in here."

Q: When you say people telling you where to go, who do you mean exactly?

A: Generally everybody that thinks that . . . well, like when I was going to school we had a special section for special education. And everybody that went by there would call us names or throw things at us and say, "We don't want to be with you; you're mentally retarded. You're crazy."

Interview II

Q: How do most people treat you when they think you are retarded?

A: Bad.

Q: What kinds of things do they do?

A: Yell at you, like when I'm with friends (unintelligible), things like that. Makes you feel about that big (gesturing to show tiny size.)

Q: Uh, huh. Does everybody do it, or does it seem like . . .

A: I'm smart enough to know the difference in people. Some of 'em treat me real nice; like everybody else. Then some of 'em don't.

Q: Does it bother you when people think you are retarded?

A: Yes, it does. I feel like I'm that big (gestured small size). It scares me a little bit.

Q: Are there things that you would like to do, but you cannot do them because somebody thinks you are retarded?

A: Yes.

Q: What are those things?

A: Well, like I've worked in a nursing home and they treat you real bad there. When they thought I was retarded.

Q: Oh really?

A: Treat me real bad. And they made me lose my job just for that.

Q: How did they treat you? What did they do exactly?

A: I tried to help them and they didn't want to be helped. And said I was retarded and all that. They were real ugly.

Q: Were these the residents at the nursing home or the other people who worked there?

A: Other people. The employees. That's the ones that you have a problem with when you go to work somewhere like that. The employees.

Q: Are there any other things that you'd like to do but can't because somebody thinks you're retarded?

A: I don't know. They start making fun of you or something. Like yesterday we were on a bus and this old lady was making fun of my husband. And I didn't like that at all. I was getting ready to tell her where to go!

Q: I imagine so. Just out of the blue?

A: Yeah, out of the blue. Look at you, and then make fun of you, and I don't think that's really right. It really does make me mad. It (unintelligible) you know, with our rights. You know we have our rights now. It's law we have our rights. If there's anything you want to, uh, through the law. I knew that long ago.

APPENDIX A

Sample Interview Schedules

Interview Schedule: Community Children

1. What is your full name?
(Probe: "What is your last name?")
2. How do you spell your name?
(Probe: "How do you spell your last/first name?")
3. Please write your name for me on this paper.
4. How old are you?
5. What is your address at home?
6. Are you usually happy or sad?
7. How many friends do you have: a lot, some, not many, or none?
8. When you're not at school, are you usually with other people?
9. Do you play any sports?
10. (If yes to "play sports", ask) What (other) sports do you play?
11. Are you usually happy?
12. Do you watch TV?
13. Do you listen to the radio or record player?
14. Do you go out to eat at restaurants or cafes?
15. Do you go to church?
16. When you're not at school, are you usually by yourself or with other people?
17. Do you know how to read books?
18. Is anybody at school teaching you about reading now?
19. Do you know how to write sentences?
20. Is anybody at school teaching you about writing now?
21. Do you know how to count money?
22. Is anybody at school teaching you about money now?
23. Are you usually sad or happy?

24. When you're not at school, are you usually by yourself?
25. Here are some ways people get to school. (Point as you say): They take the bus, somebody drives them to school in a car, they walk, or they ride to school on their bicycles. Which way do you get to school most days? Point to the picture.
(Interviewer may probe with "Most days, how do you get to school?")
26. Do you live in a house, an apartment building, a trailer house, or a duplex?
27. Are you usually sad?
28. Counting you, how many people live in your house right now?
(Interviewer probe, if necessary, with "How many is that, counting you?" or answer "Yes" if asked if we want a number.)
29. When you're not at school, are you usually with other people or by yourself?
30. If you had one wish, what would you wish for?
31. Most of the time, is it up to you to decide what time you go to bed?
32. What time does school start every day?
33. Most days, does somebody else tell you what to wear to school, or do you decide?
(Probe, if necessary, with "Most days?")
34. Is anybody at school teaching you about cooking now?
35. This picture (point) shows a boy/girl who is happy, and this one (point) shows a boy/girl who is sad. Which picture shows how you usually feel? Point to the picture.
36. Here are some pictures of different kinds of places people live in. (Point to each as you say) Here is a duplex, a house, a trailer house, and an apartment building. Which picture shows the kind of place you live in? Point to the picture.
37. Most days, how do you get to school?
(Interviewer probe, if necessary, with "Most days.")
38. What things would you really like to learn in school?
(Probe one time only with: What other things would you really like to learn in school?)
39. This picture (point) shows a boy/girl who is by himself/herself, and this one (point) shows a boy/girl who is with other people. When you're not at school, which boy/girl is most like you? Point to the picture.
40. Is anybody at school teaching you about numbers now?

41. What time do you usually go to bed at night?
(Probe, if necessary, with: "Usually.")
 42. Most days, is it up to you to decide what to wear to school, or does somebody else tell you what to wear?
(Probe, if necessary, with: "Most days.")
 43. Are you allowed to go places by yourself, or does somebody always take you?
 44. This picture (point) shows a boy/girl who is sad, and this one (point) shows a boy/girl who is happy. Which picture shows how you usually feel? Point to the picture.
 45. Most of the time, does somebody else tell you what time to go to bed?
(Probe, if necessary, with: "Most of the time.")
 46. Do you live in a duplex, a house, a trailer house, or an apartment building?
(Probe, if necessary, with: "Which one: duplex, a house, a trailer house, or an apartment?")
 47. Is anybody at school teaching you about keeping house now?
 48. Most days, do you get to school on the bus, in a car, by walking, or riding a bicycle?
(Probe, if necessary, with: "Most days." Also may have to probe: "Which one: the bus, a car, walking, or a bicycle?")
 49. Is anybody at school teaching you about getting jobs now?
 50. Are you ever allowed to go places by yourself without somebody taking you?
- Here are some pictures of different kinds of chores some people do. I want you to tell me if you do any of these chores at home.
51. Do you set the table?
(Probe, if necessary, with: "Is that a lot, some or not much?")
 52. Do you do dishes?
 53. Do you clean the floor?
 54. Do you dust furniture?
 55. Do you make beds?
(If yes:) Do you make beds a lot, some or not much?
(Probe, if necessary, with: "Is that a lot, some or not much?")
 56. Do you pick up stuff around the house?
(If yes:) Do you pick up stuff around the house a lot, some, or not much?
(Probe, if necessary, with: "Is that a lot, some, or not much?")

57. Do you cook on the stove?
58. Do you make sandwiches?
59. Do you take out the trash?
60. Do you work outside in the yard?
61. Here are some pictures of different kinds of places people live in.
(Point to each as you say:) Here is a house, an apartment building,
a trailer house, and a duplex. Which picture shows the kind of place
you live in? Point to the picture.
62. Are the neighbor kids very friendly, just so-so, or not friendly to
you?
(Interviewer probe, of necessary, with: "Which one; very friendly, or
so-so, or not friendly?")
63. This picture (point) shows a boy/girl who is with other people, and
this one (point) shows a boy/girl who is by himself/herself. When
you're not at school, which boy/girl is most like you? Point to the
picture.

Interview Schedule: Institutionalized Adults, Form A

1. What is your full name?
2. How do you spell your name?
3. Please write your name for me on this paper. (Use back of form)
4. What month is your birthday?
5. What day is your birthday?
6. What year were you born?
7. What is your address here?
8. What kind of place is this?
9. Do you have a family?
10. How often do you leave here (where you live) to go see your family?
11. How often does anybody in your family come to see you?
12. Are you usually happy or sad?
13. Are you allowed to go on dates here?
14. (If no) Why not?
15. How many friends do you have?
16. What do you and your friends usually do together?
(Probe: "What else?" until client can think of no more.)
17. What do you usually do for fun when you are by yourself?
18. Are you usually by yourself or with other people?
Tell me about the most fun thing you did last week.
19. What did you do?
20. Who was there?
21. When did it happen?
22. Where did it happen?

23. Why did you do that?
24. Who decides what chores you do?
25. Who decides how you spend your money?
- 26a. Do you watch TV?
 - b. (If yes) Do you watch TV alot, some, or not much?
27. How much do you listen to the radio, or record player: A lot, some, not much, never?
28. How much do you read books, magazines, or newspapers?
- 29a. Do you go out to the movies?
 - b. (If yes) Do you go to the movies not much, some, a lot?
30. How often do you go out to eat?
31. How often do you go to church?
- 32a. Do you go to stores?
 - b. (If yes) Do you go to stores not much, some, or a lot?
33. Do you date?
34. (If no) Why not?
35. Do you play any sports, like baseball or football?
36. (If yes to "play sports") Which ones?
37. Do you play any games indoors?
38. (If yes to "indoor games") Which ones?
39. Do you do any arts and crafts?
40. (If yes) Which ones?
41. Is anybody teaching you arts and crafts now?
42. Are you usually happy?
43. How many people sleep in your bedroom?
- 44a. Do people here yell at you or say mean things?
 - b. Do people here take things away from you?

- c. Do people here hit you?
- d. Do people here make you do things you don't want to do?
- e. Do people here bother you when you want to be alone?
- f. Do people here make fun of you?
- g. Do people here do anything else that bothers you?
- h. (If yes to "anything else that bothers you") What?
- 45a. Do people here help you when you want help?
- b. Do people here teach you things you want to learn?
- c. Do people here give you presents?
- d. Do people here hug you?
- e. Do people here talk to you when you want to talk?
- f. Do people here say nice things to you?
- g. Do people here do anything else that is nice?
- h. (If yes to "anything else nice") What?

46. Are you usually with other people?

Now I want to ask you some questions about the rules here:

- 47. Is it against the rules here to stay up late at night?
- 48. Is it against the rules here to hit people?
- 49. Is it against the rules here to call people ugly names?
- 50. Is it against the rules here to leave here without asking?
- 51. Are you usually sad?
- 52. Did you ever get punished here?
- 53. (If yes) What did you get punished for?
- 54. (If yes to "ever get punished") How were you punished?
- 55. Who makes the rules here?
- 56. Do you help make any rules here?
- 57. Which picture shows how you feel about living here. Point to the picture.

58. Why do you feel that way?
59. How do you like the food here? Not at all, not much, some, a lot.
60. Why do you feel that way?
61. How do you like the people here? A lot, some, not much, not at all.
62. Why do you feel that way?
63. Did you ever live someplace else?
64. Did you live in another place last month?
65. Did you live in another place last year?
66. Do you want to live someplace else?
67. (If yes) Why?
68. (If no) Why no?
69. Are you usually with other people or by yourself?
70. Do you have any problems?
- 71a. Do you usually have problems finding men/women friends? (same sex)
- b. Do you usually have problems finding men/women friends? (opposite sex)
- c. Do you usually have problems knowing what to do when you get sick?
- d. Do you usually have problems with anything else?
- e. (If yes) What else?
72. Who do you talk to about your problems?
(Probe: Who is that exactly?)
73. Who helps you the most when you need help?
(Probe: Who is that exactly?)
74. If you had one wish, what would you wish for?
75. Why do you wish that?
76. Do you know how to read and write?
77. Is anybody teaching you about reading and writing now?
78. Do you know how to cook?
79. Is anybody teaching you about cooking now?

80. (If yes) Who is teaching you about cooking?
(Probe: Where are you learning about cooking?)
81. Do you know how to keep house?
82. Is anybody teaching you about keeping house now?
83. Do you set the table? (show picture)
84. Do you do dishes? (show picture)
85. Do you do laundry? (show picture)
86. Do you clean the floor? (show picture)
87. Do you dust furniture? (show picture)
88. Do you make beds? (show picture)
89. Do you pick up stuff around the house/dorm? (use proper term)
90. Do you cook on the stove? (show picture)
91. Do you make sandwiches? (show picture)
92. Do you take out the trash? (show picture)
93. Who decides what chores to do?
94. Are you usually sad or happy?
95. Do you get enough food to eat?
96. Do you go to school or take classes?
97. (If yes) "What are they teaching you there?"
(Probe: "What else?" Until the client can think of no more.)
98. Do you know how to get a job?
99. Is anybody teaching you about getting jobs now?
100. (If yes) Who is teaching you about getting jobs?
(Probe: Where are you learning about getting jobs?)
101. Has anybody from vocational rehabilitation ever tried to help you?
102. (If yes) How did they help you?
103. Are you usually by yourself?
104. Do you have a job you get paid for?

105. (If no): Have you ever had a job?
106. Do you want to have a job?
107. (If yes to "Do you have a job you get paid for?): Where do you work?
108. (If yes to "have a job"): What exactly do you do at your job?
109. (If yes to "have a job"): Do you like your job?
110. (If yes to "have a job"): Do you want to keep your job?
111. Did you have a job last month?
112. Did you have a job last year?
113. Do you know how to count money?
114. Is anybody teaching you about money now?
115. (If yes) Who is teaching you about money?
(Probe: Where are you learning about money?)
116. Where do you get your money?
117. Do you get SSI?
118. Do you decide how to spend your money?
119. Could you buy some candy with \$1.00?
120. Could you buy a new bicycle with \$1.00?
121. Could you buy a new TV with \$10.00?
122. Could you buy a record album with \$10.00?

APPENDIX B

Accessing Mentally Retarded Persons for Interviewing Purposes

Accessing Mentally Retarded Persons for Interviewing Purposes

In order for any survey of a sample of the mentally retarded population to be carried out, attention to the regulations and guidelines governing participation in research projects is necessary. As indicated in Chapters 3 and 4, the procedures required in the various sites used in our interviewing project were different from one another. An interviewing project of this type depends on the cooperation of agencies and facilities to which mentally retarded persons are known. It is entirely within the discretion of those agencies and facilities to develop their own rules about access to clients, or indeed to prohibit certain types of research or any research at all. Nonetheless, there are certain guidelines that have gained common acceptance in practice, either because they are mandated or because they have been strongly advocated by key organizations. As part of our project, we solicited information about such regulations and guidelines through a letter sent to a variety of public and private organizations and agencies involved with the developmentally disabled. The following summary is intended to suggest certain key procedures and issues which must be borne in mind.

Assuming that a study has been planned and an intended population has been identified, the first step necessary is to obtain approval of the research proposal from appropriate review bodies constituted for that purpose. Institutions which receive federal research money are required to have institutional human uses committees to screen all proposals and insure that the rights of participants are adequately protected. Although such committees were originally established to review projects to be funded by the Department

of Health, Education, and Welfare, increasingly they are screening the whole range of research projects involving human subjects, whether or not they are funded externally.

Institutional human uses committees are primarily concerned with the protection of confidentiality of information, the securing of informed consent, and the protection of subjects from physical or psychological risk. The HHS regulations regarding the Protection of Human Subjects (Title 45, Part 46, revised April 1, 1977) are perhaps the most commonly followed regulations in the country because they must be adhered to if institutions are to receive HHS funding for research. In addition, studies of national scope require review of materials by the Office of Management and Budget. In our own project, for example, such a review would have been required if we had actually conducted a national survey, but was not necessary since we restricted our studies to Arkansas and Texas.

Where the cooperation of a facility or agency is required, as it inevitably would be in survey research with retarded persons, these agencies too have institutional review committees that must approve proposed research. Often more than one body is involved. In Texas, for example, a research project to be conducted at a state school for the mentally retarded must first be approved by a local committee including representation from the community. It is then forwarded to a state level committee constituted by the state Department of Mental Health and Mental Retardation for approval. In the case of Texas, then, any research to be conducted by university staff involving residents of the local state school would be approved by three bodies: the university human uses committee, and the local and state MHMR review bodies. In the case of interviewing studies, it is highly unlikely that the project would be judged to place subjects at risk, assuming that appropriate precautions are taken to insure that information provided by interviewees remains confidential

and that informed consent for participation is obtained. Nonetheless, the researcher must recognize that the process of obtaining approval can be very time consuming, taking as long as six months or so in our experience locally. Increasingly, too, local agencies and facilities are becoming hesitant to permit research which has no potential benefit for participants or for the facility staff. The benefits of participating in an interviewing study are remote (e.g., contributing to information which can be used to improve mental retardation services in the future).

Several bodies besides DHEW have developed guidelines for research which may influence local facilities and agencies and of which the mental retardation researcher should be aware. The American Association on Mental Deficiency Legislative and Social Issues Committee, for example, has developed a *Consent Handbook* (AAMD, 1977). The AAMD position is that consent is not required for the use of aggregate data (for example, from facility statistical reports), but would be required where individual subjects are involved unless the project involved anonymous interviewing about issues which are not personal in nature, in which case the subject's consent can be considered implied by his or her participation. Formal or explicit consent is recommended when the information sought might identify the respondent or anyone else, when it is likely that unsought information will accompany the release of sought information, and when the information is protected under the First Amendment. While the AAMD generally advocates compliance with DHEW guidelines for informed consent, the AAMD handbook is quite helpful in its explanation of consent as having three elements: capacity, information, and voluntariness. From the AAMD viewpoint, while DHEW emphasized information, there is a need in research with retarded persons for equal emphasis on determining the subject's capacity to understand what participation involves and insuring that the consent provided is indeed voluntary. These issues are certainly less important

in an interviewing project that involves no risk than they are in biochemical or experimental research, but the AAMD, quite rightly, emphasizes that special procedures might be necessary in view of the cognitive limitations of mentally retarded persons. Clark, Kivitz, and Rosen (1978) have discussed in detail the tendency of mentally retarded persons to comply with various requests imprudently (for example, to sign petitions without reading them or take pills without finding out whether they are harmful). Unfortunately, there are limits to the ways in which an interviewing study can be explained to potential interviewees. The soundest procedure is to give a very simple and concrete description, encourage the potential participant to ask questions, and carefully, perhaps by asking the question in a number of different ways, insure that the individual indeed wants to participate and is not simply complying without understanding. When a subject is a young child or has been declared incompetent, a substitute decision-maker (e.g., parent or guardian) can make the decision on the subject's behalf.

It should be noted that several bodies have developed positions in response to past exploitation of certain vulnerable classes of subjects, especially residents of institutions. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research has stated that such groups should no longer be overburdened by the demands of research and underbenefited by its results. Other groups, for example the American Bar Association, have taken the basic position that if research uses mentally retarded subjects it must relate to mental retardation, thereby taking a stand against using captive groups to conduct research which might be done with any number of subject pools. A DHEW proposal of 1974 sought to limit research with mentally retarded subjects to topics most likely to benefit the mentally retarded, and the American Nurses Association guidelines concur. The Joint Commission for the Accreditation of Hospitals (JCAH) recommends

that research in institutions be explained to the staff and that the research results be implemented in the institution. In the most extreme case, the Willowbrook consent judgement, a court ruled that, given a history of abuses within the facility, no research upon members of the plaintiff class would be permitted thereafter. This increased emphasis upon direct benefit evolves from an awareness of the coercive nature of institutions and the impaired decision-making capacities of the mentally retarded people who reside in them.

All documents on the use of human subjects in research emphasize the importance of maintaining a research participant's confidentiality. In particular, the problems of safeguarding against obtaining unsought information and prompting unanticipated outcomes are warned against. The confidentiality issue first crops up in identifying subjects. An interviewing study would require that agencies of facilities participate in identifying current and/or former clients. Although the regulations do not appear to be standard in this area, a hard-line position would be that only the agency or facility could make initial contact with potential subjects, and would then release to researchers names and other requested information for those who provide informed consent to participate. Indeed this was the procedure required in one of our sites. However, other agencies or facilities may be willing to give researchers access to client rosters and files on the condition that they use the information only to solicit consent for participation in the research. It is likely that agencies and facilities will become more reluctant to allow this access in the future. Whatever the regulations in a given setting, the researcher does have a clear responsibility to keep the names of potential or actual participants strictly confidential.

Interviewing studies may give rise to unanticipated problems of confidentiality. For example, the APA (1973) research participation guidelines cite the example of a study of deinstitutionalized mentally retarded persons. Risks may arise if such a study involves contacts with neighbors, relatives, and employers, who may not have previously realized that the subject was mentally retarded or institutionalized. The precaution recommended by the APA is to restrict interviews to staff who already know the subject's status and to obtain permission from the subject if anyone besides the subject is to be interviewed. There is still the risk, however, that a mentally retarded person might not fully understand the implications of having his or her employer or neighbors interviewed.

Another issue, also described in the APA document, arises when the researcher happens to acquire sensitive information, often information unrelated to the research itself. Examples especially relevant to interview studies include the discovery that a research subject abuses drugs, carries a weapon, is suicidal, or is engaging in unhealthy or destructive behavior of some sort. In such cases, researchers face the options of disclosing the information and thereby violating the subject's confidentiality or not disclosing it and possibly being responsible, in a moral if not legal sense, for subsequent harm to the subject or others. Responsibilities of the researcher to disclose or not to disclose vary from state to state, and quite obviously the ethical issues are complex. One potential solution is to include in informed consent statements an indication that the researchers might violate confidentiality if they feel a subject's welfare is endangered. In our own Center's research, a less severe but still difficult problem has at times arisen. Deinstitutionalized persons who have been interviewed have on occasion had transparent needs for help from service deliverers or have actually asked the interviewer to obtain help. Here, while it has been our policy to

comply with requests by referring the individual to appropriate local agencies, the research process itself has introduced a change in the life being studied. Otherwise, confidentiality has been strictly maintained. Indeed, clients had to be carefully convinced that participating would in no way jeopardize their community placements.

Of course researchers always have a responsibility to protect the information they collect and avoid giving it to other parties unless subjects' explicit permission has been obtained. Subjects must not be identified or identifiable in research reports or other dissemination vehicles unless they have given specific consent.

In summary, our conclusion is that regulations and guidelines for research do not pose insurmountable problems in an interviewing study, but are problematic. If a national study were to be conducted, problems would be compounded. Obtaining permission to conduct the research from a variety of agencies and facilities in a variety of settings is likely to be time-consuming and costly and certainly must be considered in project funding and staffing. In an interviewing study, risk is unlikely to be an obstacle to approval, but in some settings questions may be raised about the likely benefit to participants and participating facilities, and in all probability the researchers will have obligations to provide reports to participating facilities. The most difficult problem, we believe, is identifying eligible subjects initially. Where explanation letters and consent forms must be mailed by various facilities or agencies before names can be released to researchers, one may have difficulty obtaining agency cooperation because of the time involved in mailings and a low rate of returns is likely, making the representativeness of the final sample suspect. Efforts to obtain consent which involve personal contact, by phone or better yet face-to-face, are likely to yield far higher rates of return, but may still be prohibitively time-consuming

from an agency perspective if agency staff must make the contacts. The process is easiest if clients are physically located in a facility, more difficult if they only visit an office, and most difficult if former clients are to be included in the sample, for then the initial task of locating them may prove challenging and costly. By comparison, the tasks of obtaining consent and maintaining confidentiality once the project is underway are quite manageable.

The following materials may prove useful to researchers planning survey research, or any type of research, with mentally retarded populations:

American Association on Mental Deficiency: Consent Handbook. Washington, D.C.: AAMD, 1977 (Also statement on the use of human subjects for research).

American Psychological Association. Ethical principles in the conduct of research with human participants. Washington, D.C.: APA, 1973.

National Association for Retarded Citizens: Mental retardation research: Guidelines for the use of behavior procedures in state programs for retarded persons, and Guidelines for biomedical and pharmacological research procedures and the protection of human subjects in residential facilities for mentally retarded persons.

National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research: Protection of human subjects: Disclosure of research information under the Freedom of Information Act; and Research with institutionalized mentally infirm subjects: Recommendations.

U.S. Department of Health, Education, and Welfare: A variety of reports and regulations, especially Code of Federal Regulations, Public Welfare 45CFR 46, Protection of Human Subjects, 1977, revised edition.

APPENDIX C

Sample Explanation and Consent Forms

Texas Tech University

Research and Training Center in Mental Retardation (806) 742-3131
Training Division / (806) 742-2970

Dear Parent:

The President's Committee on Mental Retardation** want to find the best ways to get information from and about retarded people. We have asked the Research and Training Center in Mental Retardation at Texas Tech to find this out for them. The Research and Training Center would like Lubbock students to help with this project. We would like to talk to some students to see if they can understand and respond to the questions we have developed, and we must have your permission before we can do this.

The Research and Training Center will randomly pick about 60 students from those whose parents give their consent. Then, two trained interviewers will spend about 15 minutes during the school day talking to each student who was chosen. This is not a test; it is only an interview. The interview will not change your child's school programs or placement in any way.

The students will be asked some questions like "What is your birthday?" to see if they can give accurate information about themselves. They will be asked about their feelings about themselves, others, and their situations. For example, one question will be "Are you usually happy or sad?" They will also be asked to tell about activities they are involved in like sports, indoor games, arts and crafts, and chores. Some of the questions ask about services they are getting and people who help them. (For example: "Is anybody teaching you to read now?") Finally, there will be some questions about rules in the home and school and decision making. (These will be questions like: "Are you ever allowed to go places by yourself without someone taking you?")

All of this is completely anonymous. Everything your child tells the interviewers will be secret. We will not even tell anyone at the school what his/her answers are. In fact, the Research and Training Center is not really interested in what your child's answers are; we mainly want to know how well your child understands our questions. Also, if there are questions your child does not want to answer, he/she does not have to. If the interviewers are told or get the feeling that a student does not want to finish the interview, they will stop.

There may be some students who will not be able to communicate very well or at all. We still need permission for them to be part of the study because one thing the President's Committee wants to know is how many children cannot respond at all to the questions.

**The President's Committee on Mental Retardation is a committee of people appointed by the President of the United States who help the President and Congress make policies and laws for retarded citizens of our country.

Besides the student interviews, the Research and Training Center would like to spend about 20 minutes interviewing a parent or someone else who lives with the student and knows him/her very well. We will be asking you questions about your child to find out if the answers your child gave were accurate. For example, if a student answered that he sets the table "a lot," would the parent also say that their child sets the table "a lot," or would he/she say "not much?" For this parent interview, the interviewers will contact you before they come to your home or any place you would like at a time that would be best for you; even nights or weekends if necessary. There is a place for your phone number and address on the consent form. Please fill this in so they can contact you to schedule the parent interview.

Finally, the Research and Training Center would like to have your permission to use some basic information about your child from his/her school records. We would like the school to tell us your child's birthday, IQ score, and if he/she has other handicaps.

If you and your child help with this project, the Research and Training Center will be glad to give you a copy of the results. While this report will not tell you anything about your child individually, it will give you some information about how children feel about their lives and what their strengths and weaknesses in communication skills are as a group.

If you have any questions about this project, please feel free to call Carol Schoenrock at the Research and Training Center. Her phone number is 742-3134 or 742-3131, extension #38. She will be happy to answer any questions you might have about this study.

When you have decided whether or not you want your child to help with this project, please fill out the attached form, showing that you either give your permission or do not give your permission for your child to participate. Please give this form to your child to take back to school tomorrow. Thank you for your consideration.

Sincerely,

Carol K. Sigelman

Carol K. Sigelman, Ph.D.
Director of Research

Texas Tech University

Research and Training Center in Mental Retardation / (806) 742-3131
Training Division / (806) 742-2970

August 7, 1978

Dear Advocate:

The President's Committee on Mental Retardation has asked the Research and Training Center to undertake a project for them. They want us to determine the feasibility of establishing a national polling system that would enable them to quickly gain information about the needs, circumstances and attitudes of retarded citizens. By getting input from mentally retarded consumers themselves, they hope to make national policy formation more sensitive to the actual needs and opinions of these consumers.

As part of this study, we are presently trying to determine the best methods for obtaining certain types of information from and about retarded persons; in this case, adults living in the community. The Citizen Advocacy office in Austin is helping us to establish a sample of such persons to interview, and your name was given to us because you are an advocate of someone we want to include in that sample.

During the last few weeks in August, a trained interviewer from the Center will conduct personal interviews with the proteges. These interviews should take approximately 30 minutes each and will focus on residential living circumstances and on services received. We will address ourselves to the extent of the proteges' knowledge about these areas as well as their attitudes toward them. In addition to this, a telephone interview will be conducted with each person's advocate. (A personal interview can be arranged if the advocate desires.) This 15-minute interview will cover the same topics found in the protege interview, allowing us not only to determine the accuracy of the protege's responses to different types of questions, but also to obtain a different perspective. Of course, both interviews will be completely anonymous and confidential.

Before we can interview you or your protege, we need written, informed consent from both of you. Enclosed with this letter is an explanation of our project which we ask that you read to and discuss with your protege before he/she signs the consent form (also enclosed.) If your protege is not capable of giving informed consent, please discuss this with his/her guardian and let him/her sign the consent form in the space labeled GUARDIAN CONSENT. Regardless of who signs the consent form, we would appreciate your informing the guardian (or other person acting as caretaker unless the protege lives independently) that we will be contacting them or your protege to arrange a time and place to conduct the interview.

We have enclosed a self-addressed, stamped envelope for your convenience in returning the consent form. This needs to be returned whether both of you give your consent or not. The President's Committee also wants information about refusal rates. However, we hope you will help with this project, for we can all benefit from techniques to help retarded citizens express their points of view. If you or anyone else have any questions, I would be more than happy to accept a collect call at (806) 742-3131.

Sincerely,

Carol Jean Schoenrock
Project Coordinator

Enc. 2

C.3

244

EXPLANATION OF R & T CENTER PROJECT

(To be read to the protege by the advocate)

1. Some people who work at Texas Tech University in Lubbock want to come and talk to lots of people who live in Austin. If you want to, they would like to talk to you for about 10 minutes.
2. They want to ask you questions about where you live, what places you go to get help or learn things, who is teaching you things, what things you and I do together, and how you feel about where you live and the help you get.
3. They want to find out the best way to ask people these questions, and they want to find out what things are like for people who live in Austin.
4. No matter what you tell them it will not make any changes in your life. They are not trying to test or evaluate you; they want to test their questions to see if the questions are good enough to use with people all over the United States.
5. You do not have to talk to them at all if you do not want to. If you do talk to them, everything you say will be a secret. Nobody else except the person you talk to will know exactly what you say. They won't even tell me what you say. If you talk to them and there is a question you don't want to answer, you don't have to; you can just tell them you don't want to answer that.
6. They also want to talk to me and ask me some questions about you and the things I do with you and for you. Everything I say will be a secret, too. They won't tell anybody else, even you, what I tell them.
7. Before they interview you or me, they want to look at our records in the Citizen Advocacy office to get some basic information.
8. They will send a report to the Citizen Advocacy office and to some other people about the interviews they do with everyone here in Austin. But it will not say anything about you or me. It will say things like: "Such and such was a good question; most people understood it," or "This was not a good question to ask; we need to find a better way to ask it." Also, they will say things in the report like: "Most proteges said their advocate did such and such with them," and "Many of the proteges would like to learn such and such."
9. Before they interview you, they will call or come by your house to set up a time to do the interview. That way, you can tell them the best time and place for you to do it so they won't bother you when you are going to be busy.
10. If you have any questions about this project that I can't answer, we can call the people at Texas Tech and ask them before you decide if you want to help them.

C.4

245

11. When you decide, you need to sign the consent form and check one of the blanks telling them that you do want to help or that you do not want to help. Then I will send it back to them. If you check that you do want to help, they will get in touch with you in a couple of weeks to do the interview.

ADVOCATE'S NAME: _____ PROTEGE'S NAME: _____

CONSENT FORM

I understand the project that the Research and Training Center is doing. I understand that I will be interviewed and that what I say will be secret and anonymous. I know that I do not have to answer any questions I do not want to answer and that this interview will not mean any changes in my situation. I also understand that the interviewer will use information from my records at the Citizen Advocacy office.

PROTEGE:

(check one)

_____ YES, I do want to talk to the people from Texas Tech.

_____ No, I do not want to talk to the people from Texas Tech.

PROTEGE'S SIGNATURE: _____ DATE: _____

GUARDIAN CONSENT: _____ DATE: _____

ADVOCATE:

(check one)

_____ YES, I do want to talk to the people from Texas Tech.

_____ No, I do not want to talk to the people from Texas Tech.

ADVOCATE'S SIGNATURE _____ DATE: _____

XX

IF YOU HAVE AGREED TO HELP WITH THIS PROJECT, PLEASE COMPLETE THE FOLLOWING INFORMATION: PROTEGE:

Age: _____ Sex: _____ Telephone number: _____

Address (include town & zip code): _____

If we should contact someone other than the protege to set up the appointment for the interview, please give us their name (plus phone and address if different from the protege's): _____

ADVOCATE:

Telephone number: _____

Address (include town & zip code): _____

If there is any time that is generally better than another for us to contact you or your protege, please indicate below:

Advocate:

Protege:

SAMPLE INTERVIEW SCHEDULE PAGE

The following page is a reproduction of the first page of the interview schedule used with community children. Regarding the method used to code answers, the "C" notation refers to content and is where the respondent's answer is coded. Multiple C columns were used for open-ended questions in which several answers could be enumerated. The "R" variable for each question is the responsiveness code for that question. If the interviewee was unable to respond adequately to the question - for example if his answer was irrelevant - the R code would indicate that and the C code would designate missing data. This system made it easy to keypunch data directly from the interview schedule.

Coding System Used on Interview Schedules

1. ARE YOU <i>USUALLY</i> HAPPY OR SAD?		HS	C: _____:
Sad = 1	Happy = 2	HS1	R: _____:
2. ARE YOU <i>USUALLY</i> WITH OTHER PEOPLE?		WOP	C: _____:
No = 1	Yes = 2	WOP1	R: _____:

The blanks in the right-hand column permit keypunching directly from the actual interview schedules. The respondent answered the first question "sad". The "C" in the right-hand column indicates content coding, and the "sad" response is accordingly coded.

1. The "R" indicates a responsiveness code; the 8 represents a minimally appropriate response. The numbers beneath these coding blanks indicate the correct column on the cards used in keypunching.

CLIENT INTERVIEW
Consumer Study

Client Name: _____

Interviewer Name: _____

Date of Interview: _____

Research and Training Center
in Mental Retardation

Time Scheduled: _____

Texas Tech University
Lubbock, Texas

Beginning Time: _____

Ending Time: _____

COMMUNITY CHILDREN

STUDY

: 0 : 3 :
1 2

CARD

: 0 : 4 :
3 4

CASE

: : : : :
5 6 7 8

INTNUM

: : :
9 10

1. WHAT IS YOUR FULL NAME?

(Probe: "What is your last name?" Interviewer may answer questions: no middle name needed, real name; not nickname, etc.)

C : :
11

R : :
12

Incorrect = 1 Partial = 2 Correct = 3

2. HOW DO YOU SPELL YOUR NAME?

(Probe: "How do you spell your last/first name?" Int. may answer questions: first and last name, not middle name, real name, etc.)

C : :
13

R : :
14

Incorrect = 1 Partial = 2 Correct = 3

3. PLEASE WRITE YOUR NAME FOR ME ON THIS PAPER.

(Use back of form. Int. may answer questions: first and last name, not middle name, real name, etc.)

C : :
15

R : :
16

Incorrect = 1 Partial = 2 Correct = 3

CLIENT CONSENT AND INTRODUCTION

HI, MY NAME IS _____. I'M TRYING TO FIND OUT THE BEST WAY TO ASK PEOPLE QUESTIONS. A LOT OF PEOPLE HERE IN LUBBOCK ARE GOING TO HELP ME, AND I'D LIKE FOR YOU TO HELP ME, TOO. YOU DON'T HAVE TO DO THIS IF YOU DON'T WANT TO, JUST LET ME TELL YOU WHAT I WANT YOU TO DO, AND THEN YOU CAN TELL ME IF YOU WANT TO HELP. I WANT TO ASK YOU SOME QUESTIONS ABOUT YOURSELF; MOSTLY ABOUT HOW YOU FEEL AND WHAT YOU DO AT HOME AND AT SCHOOL. IT WILL TAKE ABOUT 15 MINUTES. EVERYTHING YOU TELL ME WILL BE A SECRET. I WON'T TELL ANYBODY ELSE WHAT YOU SAY. IF YOU DON'T WANT TO ANSWER A QUESTION, YOU DON'T HAVE TO. JUST TELL ME YOU DON'T WANT TO ANSWER IT. THIS IS NOT A TEST. ANY ANSWER YOU GIVE ME IS OK. I'M JUST TRYING TO FIND OUT IF PEOPLE CAN UNDERSTAND MY QUESTIONS. DO YOU WANT TO HELP ME DO THIS?

(If no, stop here. Thank client for taking time to listen.)

(If yes, continue with explanation below.)

OK, GOOD. BEFORE WE START, LET ME TELL YOU SOME THINGS. SOME OF THE QUESTIONS MAY SOUND SILLY TO YOU, BUT I HAVE TO ASK THEM ALL, EVEN IF I ALREADY KNOW THE ANSWER, I STILL HAVE TO ASK IT. I'M GOING TO BE ASKING A LOT OF THE QUESTIONS MORE THAN ONCE. THAT WON'T MEAN THAT YOU GAVE ME THE WRONG ANSWER THE FIRST TIME. THE ONLY REASON I'LL ASK SOME OF THE QUESTIONS MORE THAN ONCE IS TO FIND OUT WHICH WAY OF ASKING THEM IS THE BEST WAY. REMEMBER, TELL ME IF YOU DON'T WANT TO ANSWER A QUESTION I'LL STOP ASKING YOU.

REFERENCES

Reference Notes

Mittler, P. Comprehension of spoken language in normal and subnormal children. Paper presented to the New York Medical College Symposium, Royal Society of Medicine, London, 1970.

Shubert, O. The incidence rate of stuttering in a matched group of institutionalized mentally retarded. Paper presented at the American Association of Mental Deficiency, Chicago, 1966.

Publications Based on Study

- Budd, E.C., Sigelman, C.K., & Sigelman, L. Exploring the outer limits of response bias. Sociological Focus, 1981, 14, 297-307.
- Sigelman, C.K., Budd, E.C., Spanhel, C.L., & Schoenrock, C.J. Asking questions of retarded persons: A comparison of yes-no and either-or formats. Applied Research in Mental Retardation, 1981, 2, 347-357. (a)
- Sigelman, C.K., Budd, E.C., Spanhel, C.L., & Schoenrock, C.J. When in doubt, say yes: Acquiescence in interviews with retarded persons. Mental Retardation, 1981, 19, 53-58. (b)
- Sigelman, C.K., Budd, E.C., Winer, J.L., Schoenrock, C.J. & Martin, P.W. Evaluating alternative techniques of questioning retarded persons. American Journal of Mental Deficiency, 1982, 86, 511-518.
- Sigelman, C.K., Schoenrock, C.J., Spanhel, C.L., Hromas, S.G., Winer, J.L., Budd, E.C., & Martin, P.W. Surveying mentally retarded persons: Responsiveness and response validity in three samples. American Journal of Mental Deficiency, 1980, 84, 479-486.
- Sigelman, C.K., Schoenrock, C.J., Winer, J.L., Spanhel, C.L., Hromas, S.G., Martin, P.W., Budd, E.C., & Bensberg, G.J. Issues in interviewing mentally retarded persons: An empirical study. In R.H. Bruininks, C.E. Meyers, B.B. Sigford, & K.C. Lakin (Eds.), Deinstitutionalization and community adjustment of mentally retarded people (Monograph No. 4). Washington, D.C.: American Association on Mental Deficiency, 1981.
- Sigelman, C.K., Winer, J.L. & Schoenrock, C.J. The responsiveness of mentally retarded persons to questions. Education and Training of the Mentally Retarded, 1982, 17, 120-124.

References

- Badt, M. I. Levels of abstraction in vocabulary definitions of mentally retarded school children. American Journal of Mental Deficiency, 1958, 63, 241-246.
- Baker, B. L., Seltzer, G. B., and Seltzer, M. M. As close as possible: Community residences for retarded adults. Boston: Little, Brown and Company, 1977.
- Bangs, J. L. A clinical analysis of the articulatory defects of the feeble-minded. Journal of Speech and Hearing Disorders, 1942, 7, 343-356.
- Baroff, G. S. Mental retardation: Nature, cause, and management. Washington, DC: Hemisphere Publishing Corp., 1974.
- Bartel, N. R. The development of morphology in moderately retarded children. Education and Training of the Mentally Retarded, 1970, 5, 164-168.
- Bartel, N. R., Bryen, D., and Keehn, S. Language comprehension in the mentally retarded child. Exceptional Children, 1973, 39, 375-382.
- Bass, B. M. Authoritarianism or acquiescence? Journal of Consulting Psychology, 1955, 51, 616-623.
- Beier, E. G., Starkweather, J. A., and Lambert, M. J. Vocabulary usage of mentally retarded children. American Journal of Mental Deficiency, 1969, 73, 928-934.
- Bell, N. J. IQ as a factor in community lifestyle of previously institutionalized retardates. Mental Retardation, 1976, 14, 29-33.
- Bellugi, U. The development of interrogative structures in children's speech. In K. Riegel (Ed.), The development of language functions. Ann Arbor, MI: Language Development Program, 1965. (Report No. 8).
- Belson, W., and Duncan, J. A. A comparison of the checklist and the open response questioning systems. Applied Statistics, 1962, 11, 120-132.
- Benda, C. Mongolism and cretinism. New York: Grune and Stratton, 1946.
- Bensberg, G. J., and Sigelman, C. K. Definitions and pervallence, in L. Lloyd (Ed.), Communication, Assessment, and Intervention Strategies. Baltimore: University Park Press, 1976.
- Berg, I. A., and Rapaport, G. M. Response bias in an unstructured questionnaire. Journal of Psychology, 1954, 38, 475-481.
- Berko, J. The child's learning of English morphology. Word, 1958, 14, 150-177.

- Berry, P. B. Comprehension of possessive and present continuous sentences by nonretarded, mildly retarded, and severely retarded children. American Journal of Mental Deficiency, 1972, 76, 540-544.
- Berry, P. B., and Foxen, T. Imitation and comprehension of language in severe subnormality. Language and Speech, 1975, 18, 195-203.
- Bijou, S. W., and Werner, H. Vocabulary analysis of mentally deficient children. American Journal of Mental Deficiency, 1944, 48, 364-366.
- Bijou, S. W., and Werner, H. Language analysis in brain-injured and nonbrain-injured mentally deficient children. Journal of Genetic Psychology, 1945, 66, 239-254.
- Bilovsky, D., and Share, J. The ITPA and Down's Syndrome: An exploratory study. American Journal of Mental Deficiency, 1965, 70, 78-82.
- Birch, J. W., and Matthews, H. The hearing of mental defectives: Its measurement and characteristics. American Journal of Mental Deficiency, 1951, 55, 348-393.
- Blount, W. R. Language and the more severely retarded. American Journal of Mental Deficiency, 1968, 73, 21-29.
- Blount, W. R. A comment on language, socialization, acceptance, and the retarded. Mental Retardation, 1969, 1, 33-35.
- Bulium, N., Rynders, J., and Turnure, J. Early maternal linguistic environment of normal and Down's Syndrome language-learning children. American Journal of Mental Deficiency, 1974, 79, 52-58.
- Carroll, J. B., and White, M. N. Age of acquisition norms for 220 picturable nouns. Journal of Verbal Learning and Verbal Behavior, 1973, 12, 563-567. (a).
- Carroll, J. B., and White, M. N. Word frequency and age-of-acquisition as determiners of picture-naming latency. Quarterly Journal of Experimental Psychology, 1973, 25, 85-95. (b).
- Cohen, J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement, 1960, 20, 37-46.
- Condell, J. F. Note on the use of the Ammons Full-Range Picture Vocabulary Test with retarded children. Psychological Reports, 1959, 5, 150.
- Couch, A., and Keniston, K. Yeasayers and naysayers: Agreeing response set as a personality variable. Journal of Abnormal and Social Psychology, 1960, 60, 151-174.
- Cromer, R. F. Receptive language in the mentally retarded: Processes and diagnostic distinctions. In R. L. Schiefelbusch & L. L. Lloyd, (Eds.), Language perspectives: Acquisition, Retardation, and Intervention. Baltimore: University Park Press, 1974.

- Cronbach, L. J. Response sets and test validity. Educational and Psychological Measurement, 1946, 6, 475-494.
- Cronbach, L. J. Studies of acquiescence as a factor in the true-false test. Journal of Educational Psychology, 1942, 33, 401-415.
- Dever, R. B. A comparison of the results of a revised version of Berko's Test of Morphology with the free speech of mentally retarded children. Journal of Speech and Hearing Research, 1972, 15, 169-198.
- Dever, R. B., and Gardner, W. I. Performance of normal and retarded boys on Berko's Test of Morphology. Language and Speech, 1970, 13, 162-181.
- DeVart, M. H. Language comprehension processes of mentally retarded children. American Journal of Mental Deficiency, 1979, 84, 177-183.
- Edgerton, R. B. The cloak of competence: Stigma in the lives of the mentally retarded. Berkeley: University of California Press, 1967.
- Edwards, A. L. The social desirability variable in personality assessment and research. New York: The Dryden Press, 1957.
- Elias, S. F., Sigelman, C. K., and Danker-Brown, P. Interview behaviors of mentally retarded adults as predictors of employability. Journal of Applied Psychology, 1980, 65, 67-73.
- Ervin-Tripp, S. Discourse agreement: How children answer questions. In J.R. Hayes (Ed.). Cognition and the development of language. New York: Wiley, 1970.
- Fraser, C., Bellugi, V., and Brown, R. Control of grammar in imitation, comprehension, and production. Journal of Verbal Learning and Verbal Behavior, 1963, 2(2), 121-135.
- Gallagher, J. J. A comparison of brain injured/non-brain injured mentally retarded children on several psychological variables. Monographs of the Society for Research in Child Development, 1957, 22, (No. 65).
- Gallagher, J. W. Short-term recall of sentences in normal and retarded children. American Journal of Mental Deficiency, 1969, 74, 57-61.
- Gerjuoy, I. and Winters, J. J. Jr. Lateral preference for identical geometric forms: II. Retardates. Perception & Psychophysics, 1966, 1, 104-106.
- Goda, S., and Griffith, B. C. Spoken language of adolescent retardates and its relation to intelligence, age, and anxiety. Child Development, 1962, 33, 489-498.
- Gollay, E., Freedman, R., Wyngaarden, M., and Kurtz, N. R. Coming back. The community experiences of deinstitutionalized mentally retarded people. Cambridge, MA: Abt Books, 1978.

- Goodstein, H. A. Performance of mentally handicapped and average-IQ children on two modified cloze tasks for oral language. American Journal of Mental Deficiency, 1970, 75, 290-297.
- Gotttsleben, R. J. The incidence of stuttering in a group of mongoloids. Training School Bulletin, 1955, 51, 208-218.
- Gozali, J., and Bialer, I. Children's locus of control scale: Independence from response set bias among retardates. American Journal of Mental Deficiency, 1968, 72, 622-625.
- Graham, J. T., and Graham, L. W. Language behavior of the mentally retarded: Syntactic characteristics. American Journal of Mental Deficiency, 1971 75, 623-629.
- Greenfield, P. M., and Smith, J. H. The structure of communication in early language development. New York: Academic Press, 1976.
- Griffith, B. C., and Spitz, H. Some relationships between abstraction and word meaning in retarded adolescents. American Journal of Mental Deficiency, 1958, 63, 247-251.
- Guess, D. A functional analysis of receptive language and productive speech: acquisition of the plural morpheme. Journal of Applied Behavior Analysis, 1969, 2, 55-64.
- Guess, D. and Baer, D. M. An analysis of individual differences in generalization between receptive and productive language in retarded children. Journal of Applied Behavior Analysis, 1973, 6, 311-329.
- Gutmann, A. J. and Rondal, G. A. Verbal operants in mother's speech to non-retarded and Down's Syndrome children matched for linguistic level. American Journal of Mental Deficiency, 1979, 83, 446-452.
- Hare, A. P. Interview responses: Personality or conformity? Public Opinion Quarterly, 1960, 24, 679-685.
- Hogan, D. D. Errors in computation of hearing loss in studies of large populations. Mental Retardation, 1973, 11, 15-17.
- Hooper, R. Communicative development and children's responses to questions. Speech Monographs, 1971, 38(1), 1-9.
- Irwin, O. C. The development status of speech sounds of ten feeble-minded children. Child Development, 1942, 13, 29-39.
- Jenkins, J. Psychology in business and industry. New York: John Wiley and Sons, 1935.
- Karlin, I. W., and Strazzula, M. Speech and language problems of mentally deficient children. Journal of Speech and Hearing Disorders, 1952, 17, 286-294.

- Katz, D. Do interviewers bias poll results? Public Opinion Quarterly, 1942, 6, 248-268.
- Keane, V. E. The incidence of speech and language problems in the mentally retarded. Mental Retardation, 1972, 10, 3-8.
- Knudsen, D. D., Pope, H., and Irish, D. P. Response differences to questions on sexual standards: An interview-questionnaire comparison, Public Opinion Quarterly, 1967, 31, 290-297.
- Kodman, F. The incidence of hearing loss in mentally retarded children. American Journal of Mental Deficiency, 1958, 62, 675-678.
- Kodman, F., Powers, T. R., Philip, P., and Weller, G. M. An investigation of hearing loss in mentally retarded children and adults. American Journal of Mental Deficiency, 1958, 63, 460-463.
- Lachman, R. Uncertain effects of time to access the internal lexicon. Journal of Experimental Psychology, 1973, 99, 199-208.
- Lachman, R., Schaffer, J. B., and Hennrikus, D. Language and cognition: Effects of stimulus codability, name-word frequency, and age-of-acquisition on lexical reaction time. Journal of Verbal Learning and Verbal Behavior, 1974, 13, 613-625.
- Lackner, J. R. A developmental study of language behavior in retarded children. In D. M. Morehead & A. E. Morehead (Eds.), Normal and deficient child language. Baltimore: University Park Press, 1976.
- Lamberts, F., and Weener, P. D. TMR children's competence in processing negation. American Journal of Mental Deficiency, 1976, 81, 181-186.
- Layton, T. L., and Sharifi, H. Meaning and structure of Down's Syndrome and nonretarded children's spontaneous speech. American Journal of Mental Deficiency, 1978, 83, 439-455.
- Lenneberg, E. H. Biological foundations of language. New York: Wiley, 1967.
- Lenneberg, E. H., Nichols, I. A., and Rosenberger, E. F. Primitive stages of language development in mongolism. In D.M. Rioch & E.A. Weinstein (Eds.), Disorders of Communication. (Research publications of the Association for Research in Nervous and Mental Disease, Vol. XLII). Baltimore: Williams & Wilkins, 1964.
- Lenski, G. E., and Leggett, J. C. Caste, class and deference in the research interview. American Journal of Sociology, 1960, 65, 463-467.
- Leonard, L. B., Cole, B., and Steckol, K. F. Lexical usage of retarded children: An examination of informativeness. American Journal of Mental Deficiency, 1979, 84, 49-54.

- Lieber, C. W., and Spitz, H. H. Inference of word meaning from syntax structure by normal and retarded adolescents. Journal of Psychology, 1976, 93, 3-12.
- Little, N. F., and Williams, H. M. An analytic scale of language achievement. University of Iowa Studies in Child Welfare, 1937, 13(2).
- Lloyd, L. L. Audiologic aspects of mental retardation. In N. R. Ellis (Ed.), International Review of Research in Mental Retardation (Vol.4). New York: Academic Press, 1970
- Lloyd L. L., and Cox, B. P. Programming for the audiologic aspects of mental retardation. Mental Retardation, 1972, 10, 22-26.
- Lloyd, L. L., and Reid, M. J. The incidence of hearing impairment in an institutionalized mentally retarded population. American Journal of Mental Deficiency, 1967, 71, 746-763.
- Lovell, K., and Bradbury, B. The learning of English morphology in educationally subnormal special school children. American Journal of Mental Deficiency, 1967, 71, 609-615.
- Lozar, B., Wepman, J. M., and Hass, W. Lexical usage of mentally retarded and nonretarded children. American Journal of Mental Deficiency, 1972, 76, 534-539.
- Lyle, J. G. The effect of an institutional environment upon the verbal development of imbecile children: I. Verbal intelligence. Journal of Mental Deficiency Research, 1959, 3, 122-128.
- Lyle, J. G. The effect of an institutional environment upon the verbal development of imbecile children: II. Speech and language. Journal of Mental Deficiency Research, 1960, 4, 1-13. (a)
- Lyle, J. G. The effect of an institutional environment upon the verbal development of imbecile children: III. The Brooklands residential family unit. Journal of Mental Deficiency Research, 1960, 4, 24-23. (b)
- Lyle, J. G. Some factors affecting the speech development of imbecile children in an institution. Journal of Child Psychology and Psychiatry, 1960, 1, 121-129. (c)
- Lyle, J. G. A comparison of the verbal intelligence of normal and imbecile children. Journal of Genetic Psychology, 1961, 99, 277-324. (a)
- Lyle, J. G. Comparison of the language of normal and imbecile children. Journal of Mental Deficiency Research, 1961, 5, 40-51. (b)
- Marshall, N., Hegrenes, J., and Goldstein, S. Verbal interactions: Mothers and their retarded children versus mothers and their nonretarded children. American Journal of Mental Deficiency, 1973, 77, 415-419.

- Matthews, J. Communication disorders in the mentally retarded. In L. E. Travis (Ed.), Handbook of speech pathology and audiology. New York: Appleton-Century-Crofts, 1971.
- McCarthy, D. Language development in children. In L. Carmichael (Ed.), Manual of child psychology. New York: John Wiley & Sons, 1954.
- McNutt, J. C., and Leri, S.M. Language differences between institutionalized and noninstitutionalized retarded children. American Journal of Mental Deficiency, 1979, 83, 339-345.
- Mein, R. A study of the oral vocabularies of severely subnormal patients. II. Grammatical analysis of speech samples. Journal of Mental Deficiency Research, 1961, 5, 52-59.
- Mein, R. Use of the Peabody Picture Vocabulary Test with severely subnormal patients. American Journal of Mental Deficiency, 1962, 67, 269-273.
- Mein, R., and O'Connor, N. A study of the oral vocabularies of severely subnormal patients. Journal of Mental Deficiency Research, 1960, 4, 130-143.
- Miller, M. A., Cuvo, A. J., and Borakove, L. Teaching naming of coin values: Comprehension before production versus production alone. Journal of Applied Behavior Analysis, 1977, 10, 735-736.
- Montague, J. C., Hutchinson, E. C., and Matson, E. A comparative computer content analysis of the verbal behavior of the institutionalized and noninstitutionalized retarded children. Journal of Speech and Hearing Research, 1975, 18, 43-57.
- Nirje, B. The right to self-determination. In W. Wolfensberger (Ed.), The Principle of Normalization in Human Services. Toronto: National Institute of Mental Retardation, 1972.
- Papinia, N. A. A qualitative analysis of vocabulary responses of institutionalized mentally retarded children. Journal of Clinical Psychology, 1954, 10, 361-365.
- Piaget, J. [The child's conception of time.] (A. J. Pomeran, trans.). New York: Basic Books, 1970. (Originally published, 1927).
- Reynolds, W. M., and Reynolds, S. Prevalence of speech and hearing impairment of noninstitutionalized mentally retarded adults. American Journal of Mental Deficiency, 1979, 84, 62-66.
- Rigrodsky, S., and Goda, S. Language behavior of a group of non-speaking brain damaged mentally retarded children. Training School Bulletin, 1951, 58, 52-60.

- Rigrodsky, S., Prunty, F., and Glovsky, L. A study of the incidence, types, and associated etiologies of hearing loss in an institutionalized mentally retarded population. Training School Bulletin, 1961, 58, 30-44.
- Rittmanic, P. A. Hearing rehabilitation for the institutionalized mentally retarded. American Journal of Mental Deficiency, 1959, 63, 778-783.
- Rohr, A., and Burr, D. B. Etiological differences in patterns of psycholinguistic development of children of IQ 30 to 60. American Journal of Mental Deficiency, 1978, 82, 549-553.
- Rorer, L. G. The great response-style myth. Psychological Bulletin, 1965, 63, 129-156.
- Rosen, M., Clark, G. R., and Kivitz, M. S. Habilitation of the Handicapped. Baltimore: University Park Press, 1977.
- Rothenberg, B. Conservation of number among four and five-year-old children: Some methodological considerations. Child Development, 1969, 40, 332-406.
- Rugg, D. Experiments in wording questions: II. Public Opinion Quarterly, 1941, 5, 91-92.
- Schaeffer, M. L., and Shearer, W. M. A survey of mentally retarded stutterers. Mental Retardation, 1968, 6, 44-45.
- Schlanger, B. B. Speech measurements of institutionalized mentally handicapped children. American Journal of Mental Deficiency, 1953, 58, 114-122. (a)
- Schlanger, B. B. Speech examination of a group of institutionalized mentally handicapped children. Journal of Speech and Hearing Disorders, 1953, 18, 339-349. (b)
- Schlanger, B. B. Environmental influences on the verbal output of mentally retarded children. Journal of Speech and Hearing Disorders, 1954, 19, 339-343.
- Schlanger, B. B. Results of varying presentations to brain damaged children of an auditory work discrimination test. American Journal of Mental Deficiency, 1958, 63, 464-468.
- Schlanger, B. B., and Gottsleben, R. H. Testing the hearing of the mentally retarded. Journal of Speech and Hearing Disorders, 1956, 21, 487-493.
- Schlanger, B. B., and Gottsleben, R. H. Analysis of speech defects among the institutionalized mentally retarded. Journal of Speech and Hearing Disorders, 1957, 22, 98-103.
- Schuman, H., and Converse, J. M. The effects of black and white interviewers on white respondents in 1968. Public Opinion Quarterly, 1971, 35, 44-68.

- Schuman, H., and Presser, S. Question wording as an independent variable in survey analysis. Sociological Methods & Research, 1977, 6, 151-170.
- Semmel, M. I., Barritt, L. S., and Bennett, S.W. Performance of EMR and nonretarded children on a modified cloze task. American Journal of Mental Deficiency, 1970, 74, 681-688.
- Semmel, M. I., Barritt, L. S., Bennett, S. W., and Perfectti, C. A. A grammatical analysis of word association of educable mentally retarded and normal children. American Journal of Mental Deficiency, 1968, 72, 567-576.
- Semmel, M. I., and Dolley, D. G. Comprehension and imitation of sentences of Down's Syndrome children as a function of transformational complexity. American Journal of Mental Deficiency, 1971, 75, 739-745.
- Siegenthaler, B. M., and Krzywicki, D. F. Incidence and patterns of hearing loss among an adult mentally retarded population. American Journal of Mental Deficiency, 1959, 64, 444-449.
- Sievers, D., and Essa, S. H. Language development in institutionalized and community mentally retarded children. American Journal of Mental Deficiency, 1961, 66, 413-420.
- Sigelman, C. K., and Werder, P. R. The communication skills of the mentally retarded: A new analysis. Journal of Developmental Disabilities, 1975, 1, 9-16.
- Sirkin, J., and Lyons, W. A study of speech defects in mental deficiency. American Journal of Mental Deficiency, 1941, 46, 74-80.
- Skeels, H. M., Updegraff, R., Wellman, B., and Williams, H. M. A study of environmental stimulation: An orphanage preschool project. University of Iowa Studies in Child Welfare, 1938, 15(4).
- Sloan, W., and Bensberg, G. J. An exploratory study of the Full-Range Picture Vocabulary Test with mental defectives. American Journal of Mental Deficiency, 1954, 58, 481-485.
- Sloan, W., and Cutts, R. A. Test patterns of mental defectives on the revised Stanford-Binet Scale. American Journal of Mental Deficiency, 1947, 51, 394-396.
- Slobin, D. I. Grammatical transformations and sentence comprehension in childhood and adulthood. Journal of Verbal Learning and Verbal Behavior, 1966, 5, 219-227.
- Smith, M. E. The influence of age, sex, and situation on the frequency, form, and function of questions asked by preschool children. Child Development, 1933, 4, 201-213.

- Spitz, R. A. Hospitalism: An inquiry into the genesis of psychiatric conditions in early childhood. The Psychoanalytic Study of the Child (Vol. I), New York: International Universities Press, 1945, 53-74.
- Spitz, R. A. Hospitalism: A follow-up report on investigation described in Volume I, 1945. The Psychoanalytic Study of the Child (Vol. II), New York: International Universities Press, 1946, 113-117.
- Spradlin, J. E. Language and communication of mental defectives. In N. R. Ellis (Ed.), Handbook of Mental Deficiency. New York: McGraw-Hill, 1963.
- Spreen, O. Language functions in mental retardation: A review. Language, development, types of retardation, and intelligence level. American Journal of Mental Deficiency, 1965, 69, 482-494.
- Stanovich, K. E., and Stanovich, P. J. Speaking for themselves: A bibliography of writings by mentally handicapped individuals. Mental Retardation, 1979, 17(2), 83-86.
- Sudman, S., and Bradburn, N. M. Response effects in surveys: A review and synthesis. Chicago: Aldine, 1974.
- Thorndike, R. L., Hagen, E., and Kemper, R. A. Normative data obtained in house-to-house administration of a psychosomatic inventory. Journal of Consulting Psychology, 1952, 16, 257-260.
- Tyack, D., and Ingram, D. Children's production and comprehension of questions. Journal of Child Language, 1977, 4, 211-224.
- Vasta, R., and Teitelbaum, M. Reception training effects on the production of modeled language constructions. Journal of Experimental Child Psychology, 1976, 22, 67-72.
- Walker, H. J., Roodin, P. A., and Lamb, M. J. Relationship between linguistic performance and memory deficits in retarded children. American Journal of Mental Deficiency, 1975, 79, 545-552.
- Webb, C. E., and Kinde, S. Speech, language, and hearing of the mentally retarded. In A. A. Baumeister (Ed.), Mental Retardation: Appraisal, Education, & Rehabilitation. Chicago: Aldine Publishing Co., 1967.
- Wells, W. D. How chronic overclaimers distort survey findings. Journal of Advertising Research, 1963, 3, 8-18.
- West, R., Kennedy, L., and Carr, A. The rehabilitation of speech (2nd Ed.). New York: Harper, 1947.
- Whitehurst, G. Comprehension, selective imitation, and the CIP hypothesis. Journal of Experimental Child Psychology, 1977, 23, 23-38.
- Winters, J. J., and Bronckow, M. A. Development of language in normal and retarded persons. Psychological Reports, 1975, 37, 391-402.

Winters, J. J., and Cundari, L. Speed of retrieving information from the lexicon of mentally retarded adolescents. American Journal of Mental Deficiency, 1979, 83, 566-570.

Yoder, D. E., and Miller, J. F. What we may know and what we can do: Input toward a system. In J. E. McClean, D. E. Yoder, & R. L. Schiefelbusch (Eds.), Language Intervention with the Retarded: Developing Strategies. Baltimore: University Park Press, 1972.